New opportunities for agricultural extension services:

Mainstreaming large-scale farmer participation through modern ICT

Dissertation

zur Erlangung des

akademischen Grades

Doctor rerum agriculturarum (Dr. rer. agr.)

eingereicht an der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

Jonathan Steinke, M.Sc.

Wissenschaftliche Supervision

Dr. Jacob van Etten

GutachterInnen der Dissertation

1. Prof. Dr. Wolfgang Bokelmann
2. Dr. habil. Stefan Sieber
3. Prof. Dr. Regina Birner

Tag der mündlichen Prüfung

4. Dezember 2019

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

Prof. Dr. Bernhard Grimm

New opportunities for agricultural extension services:

Mainstreaming large-scale farmer participation through modern ICT

PhD thesis submitted by Jonathan Steinke, M.Sc.

Table of contents

i Summary

Smallholder farmers across the Global South increasingly need to adapt their farming activities to fast-paced changes, for example, in climate, policy and markets. In many places, public and private agricultural extension services support technological change through trainings and the dissemination of information. The effectiveness of extant extension (advisory) methodologies is, however, challenged by the difficulty of reaching a large and growing clientele with highly diverse information needs. In recent years, the increasing penetration of modern information and communication technology (ICT) has created new opportunities for disseminating agricultural information more widely. In addition, modern ICT may allow harnessing the existing heterogeneity of farmers and farms in a positive way. Through digital communication, large numbers of farmers can be involved not only as recipients of advice, but also in the creation of knowledge and information. By collecting well-defined data inputs from farmers through digital channels, and processing these data in systematic ways, agricultural advisory services can potentially improve their overall performance towards a large and heterogeneous clientele.

This dissertation intends to explore these emerging socio-technological opportunities. Through three proof-of-concept studies, it delivers empirical evidence on the feasibility of different ways of employing modern ICT to harness large-scale farmer participation in agricultural extension. Subsequently, it discusses potential practical implications for the ability of extension services to serve large numbers of farmers, working in heterogeneous conditions, with individually adequate advice. The dissertation follows a three-pronged approach. It focuses on three selected, but common shortcomings of agricultural extension, all of which are due to the inherent scale and complexity of the smallholder farming context that needs to be served. To each shortcoming, one research paper explores a novel concept of enabling large-scale farmer participation through modern ICT, as a potential solution.

The first deficiency considered here is that agricultural extension often treats farmers as passive recipients of information, and rarely involves them in the generation of knowledge. Given the general unpredictability that characterizes smallholder farming, however, farmers may benefit strongly from developing their heuristic and observation skills. In recent 'citizen science' projects, large numbers of farmers contribute to agricultural research by individually carrying out small on-farm experiments. Digital channels allow rapid, massive, and cost-effective collection of observations ('crowdsourcing'), as well as feedback of research results to the farmer participants. In order to assess the

usefulness of a novel agricultural citizen science approach for crop variety selection, the first research paper studies the accuracy of farmer-generated data. It then explores implications for carrying out large-scale crowdsourced variety selection trials. In experimental observation exercises with 35 farmers in Honduras, it was found that farmers make observations that fully or almost fully match an agronomist's observation in at least 77 % of cases. Although there was some disagreement among observing farmers, all in all, simulations showed that the observed level of agreement was sufficient to make valid statements on varietal quality with realistic numbers of contributors, such as 200 farmers. This result may justify crowdsourcing farmer observations as a data source in agricultural research, which would regularly expose farmers to new germplasm, strengthening their observation skills.

The second deficiency is the frequent limited context-adaptation of agricultural advisory contents. Agricultural technologies and innovations promoted by extension services sometimes show unexpected disbenefits under local farmer management, resulting in weak adoption or strong dis-adoption. The 'Positive Deviance approach' is a research methodology that acknowledges the diversity and significance of local context, and helps to identify innovative practices that are viable in local context. The second research paper adapts this approach to the context of agricultural development. It presents evidence on the feasibility of using survey data, which can be crowdsourced through digital channels, for identifying 'positive deviant' farming households. These household achieve surprisingly strong performance regarding multiple household objectives in spite of trade-offs and resource limitations. Follow-up visits with 15 'positive deviants' in Tanzania revealed 14 behaviors that plausibly contributed to superior performance, and which could be promoted to other households. The study demonstrates how the Positive Deviance approach can help extension services to rapidly identify promising practices that work in local context.

The third deficiency is the frequent weak attention to household diversity in the delivery of agricultural advice. Working on tight budgets and with little staff, extension services often cannot disaggregate generic advisory messages in response to the specific conditions of individual farms. Existing digital advisory applications allow farmers to select contents via two-way communication interfaces, such as USSD or IVR, but menus can be long and tedious. To be useful in practice, such applications should be able to further prioritize advisory contents based on very little information about the user. Therefore, the third research paper explores the feasibility of mobile phone-mediated two-way communication between farmers and an automated advisory application that prioritizes advisory messages in a household-specific way. In a triple replication of the experimental design, between 43 and 98 farmers from Ethiopia, Kenya, and Tanzania first answered a quantitative household survey, and then expressed individual preferences for different advisory contents. Using a modelling approach, it was found that with farmers answering just 5 to 10 questions through an ICT interface, household-specific prioritizations can be generated that increase the individual fit of delivered advice and reduce the risk of delivering irrelevant information. This suggests that household-specific targeting of agricultural advice based on two-way communication is feasible in practice. Because further questions can be asked each time a farmer accesses the service, targeting can continuously improve through learning algorithms that improve the underlying model.

This dissertation contributes to a growing body of scientific evidence on how agricultural extension services in the Global South can address the interconnected challenges of scale and complexity in smallholder farming context through increased methodological pluralism, greater farmer participation, and efficient, systematic use of digital media. Based on the presented evidence, the dissertation concludes with a discussion of further needs and challenges for truly mainstreaming these novel concepts in agricultural advisory services. Increased use of ICT may not only imply new channels of information dissemination. The possible diversification of advice and an increased reliance on farmer-generated data inputs also requires institutional change within extension providers. This may include, for example, reducing managerial top-down decision-making in favor of data-driven, flexible decision-making at all levels, including by field agents. Moreover, the adoption of digital communication by farmers and agricultural advisors will likely be facilitated by integrating and concentrating multiple information and communication services within well-designed digital applications. This thesis suggests how the three independent concepts studied here could be integrated into a single digital service for agricultural advisory. In the long run, policy-makers can facilitate such synergies and cooperation between digital developers and agricultural extension through fiscal incentives and the provision of local 'tech hubs'.

ii Zusammenfassung

Kleinbäuerinnen und Kleinbauern im Globalen Süden sind angesichts rasanter Veränderungen zunehmend gefordert, ihre landwirtschaftlichen Aktivitäten anzupassen, etwa an sich ändernde klimatische, regulatorische und marktwirtschaftliche Rahmenbedingungen. Vielerorts unterstützen staatliche und private landwirtschaftliche Beratungsdienste (*extension services*) den technologischen Wandel durch Schulungen und die Verbreitung von Informationen. Aufgrund der Schwierigkeit, eine große und wachsende landwirtschaftliche Bevölkerung mit heterogenen Informationsbedürfnissen adäquat zu erreichen, stößt die Effektivität derzeit verwendeter Methoden allerdings an Grenzen. Die zunehmende Verbreitung moderner Informations- und Kommunikationstechnologien (ICT, gemäß dem englischsprachigen Akronym) hat in jüngster Zeit neue Möglichkeiten geschaffen, Information weitreichend zu übertragen. Darüber hinaus bieten moderne ICT aber auch Chancen, die Heterogenität innerhalb der Zielgruppe zu verwerten. Digitale Kommunikation erlaubt es, einer großen Zahl von Bäuerinnen und Bauern Informationen zu liefern, aber auch, sie in der Erzeugung von Wissen und Information einzubinden. Über digitale Kanäle könnten landwirtschaftliche Beratungsdienste Daten-Inputs von Bäuerinnen und Bauern erheben, sie systematisch verarbeiten, und auf dieser Grundlage ihre Dienstleistung für eine große und heterogene Zielgruppe verbessern.

Diese Dissertationsschrift zielt darauf ab, die entstehenden technologisch-sozialen Möglichkeiten zu erkunden. Drei Machbarkeitsstudien präsentieren empirische Erkenntnisse zur Umsetzbarkeit verschiedener Strategien zur Einbindung großer Zahlen von Bäuerinnen und Bauern in der landwirtschaftlichen Beratung mittels moderner ICT. Anschließend folgt eine Diskussion der potenziellen praktischen Auswirkungen für die Fähigkeit von Beratungsdiensten, große Zahlen von Bäuerinnen und Bauern, die unter heterogenen Bedingungen arbeiten, mit individuell angemessenen Auskünften zu adressieren. Die Dissertation folgt einem dreigliedrigen Ansatz. Sie behandelt drei ausgewählte, aber häufige Unzulänglichkeiten in der Praxis der landwirtschaftlichen Beratung. Diese Herausforderungen führen alle auf das Ausmaß und die inhärente Komplexität des kleinbäuerlichen Kontexts zurück. Zu jeder dieser Unzulänglichkeiten untersucht eine veröffentlichte Forschungsarbeit ein neues Konzept zur Partizipation großer Zahlen von Bäuerinnen und Bauern mittels moderner ICT, als mögliche Lösung.

Das erste Problem, das hier betrachtet wird, besteht darin, dass landwirtschaftliche Beratungsdienste häufig Bäuerinnen und Bauern als passive EmpfängerInnen von Information behandeln, sie aber selten in die Erzeugung von Wissen einbeziehen. Angesichts der

allgemeinen Unvorhersehbarkeit, die die kleinbäuerliche Produktion kennzeichnet, kann die bäuerliche Praxis jedoch stark von der Förderung heuristischer Fähigkeiten und Beobachtungsgaben profitieren. In neuen Projekten der ‚Bürgerwissenschaft' (*Citizen Science*) leisten zahlreiche Bäuerinnen und Bauern einen Beitrag zur Agrarforschung, indem sie auf ihren eigenen Betrieben individuell kleine Experimente durchführen. Digitale Kanäle ermöglichen dann das schnelle und kostengünstige Erheben von vielen Beobachtungen (‚crowdsourcing') ebenso wie die Rückmeldung der analysierten Ergebnisse an die teilnehmenden Bäuerinnen und Bauern. Um die Nutzbarkeit eines neuartigen landwirtschaftlichen Citizen Science-Ansatzes zur Sorten-Selektion zu bewerten, untersucht die erste Forschungsarbeit die Genauigkeit der von den Bäuerinnen und Bauern generierten Daten. Anschließend werden Implikationen für die Durchführung von großangelegten Sortenauswahl-Experimenten mittels Crowdsourcing untersucht. Hierfür wurden Citizen Science-Datenerhebungen mit 35 Bäuerinnen und Bauern in Honduras experimentell nachgestellt. Es zeigte sich, dass die Beobachtungen der Bäuerinnen und Bauern in mindestens 77 % der Fälle mit jenen eines Agronomen vollständig oder fast vollständig übereinstimmten. Trotz eines gewissen Maßes an Nichtübereinstimmung unter den Bäuerinnen und Bauern zeigten Simulationen, dass das beobachtete Maß an Übereinstimmung insgesamt ausreicht, um mit einer realistischen Anzahl von Mitwirkenden – zum Beispiel 200 Bäuerinnen und Bauern – valide Aussagen zur Qualität der untersuchten Sorten zu treffen. Mit diesem Ergebnis lässt sich das Crowdsourcing von bäuerlichen Beobachtungen als Datenquelle in der Agrarforschung rechtfertigen, was die TeilnehmerInnen regelmäßig neuem Sortenmaterial aussetzen, und dadurch ihre Beobachtungsgaben stärken würde.

Das zweite Problem ist die teils eingeschränkte Eignung der landwirtschaftlichen Beratungsinhalte unter den Bedingungen des Zielkontexts. Unter lokalen Bedingungen können die von Beratungsdiensten empfohlenen landwirtschaftliche Technologien unvorhergesehene Nachteile entfalten, was zu geringer oder nur kurzfristiger Übernahme seitens der Bäuerinnen und Bauern führt. Der ‚Positive Deviance'-Ansatz ist eine Forschungsmethode, die die Vielfältigkeit sowie die Bedeutung des lokalen Kontexts anerkennt und dabei hilft, lokal umsetzbare, innovative Praktiken zu identifizieren. Die zweite Forschungsarbeit adaptiert diesen Ansatz für den Kontext der landwirtschaftlichen Entwicklung. Sie liefert Erkenntnisse zur Durchführbarkeit der Methode, die unter Verwendung von quantitativen Haushaltsdaten, welche über digitale Kanäle erhoben werden können, sogenannte ‚positive deviants' identifiziert. Diese sind Haushalte, die trotz begrenzter Ressourcen sowie Zielkonflikten eine überraschend starke Leistung hinsichtlich mehrerer Entwicklungsziele erreichen. Durch Besuche bei 15 solcher ‚positive deviants' in Tansania wurden 14 Praktiken identifiziert, die mit einiger Plausibilität zur besseren Zielerreichung beitrugen, und die anderen Haushalten empfohlen werden könnten. Die Studie zeigt, wie der Positive Deviance-Ansatz Beratungsdiensten helfen kann, auf effiziente Weise vielversprechende Praktiken zu identifizieren, die im lokalen Kontext funktionieren.

Als drittes Problem behandelt die vorliegende Dissertationsschrift die häufig unzureichende Berücksichtigung der Vielfältigkeit kleinbäuerlicher Haushalte bei der Bereitstellung von landwirtschaftlichen Beratungsleistungen. Angesichts eng bemessener Budgets und knapper Personalbestände sind Beratungsdienste oft nicht in der Lage, Beratungsinhalte auf die spezifischen Bedingungen einzelner Betriebe anzupassen. Bestehende digitale Systeme für die Agrarberatung ermöglichen es Landwirten zwar bereits, individuell angepasste Inhalte über Kommunikationsschnittstellen (*two-way communication*), etwa USSD oder IVR, auszuwählen. Die Menüs sind allerdings häufig langwierig und mühsam zu bedienen. Um in der Praxis nützlich zu sein, sollten digitale Systeme fähig sein, Beratungsinhalte auf Basis sehr weniger Informationen über die Nutzerin individuell zu priorisieren. Die dritte Forschungsarbeit untersucht daher die Umsetzbarkeit von Mobiltelefon-gestützter Zwei-Wege-Kommunikation zwischen Bäuerinnen und Bauern einerseits sowie einem automatisierten Beratungssystem andererseits, welches Beratungsinhalte für individuelle Haushalte priorisiert. In einer dreifachen Replikation des Versuchsaufbaus wurden zunächst quantitative Haushaltsdaten von 48 bis 98 Bäuerinnen und Bauern in Äthiopien, Kenia und Tansania gesammelt. Im Anschluss äußerten die Befragten individuelle Präferenzen für verschiedene Beratungsinhalte. Eine ökonometrische Modellierung zeigte, dass es bereits auf Basis von 5 bis 10 Fragen, die via ICT beantwortet werden können, möglich ist, individuelle Priorisierungen von Beratungsinhalten zu generieren, welche die individuelle Passung der bereitgestellten Inhalte spürbar erhöhen sowie das Risiko der Bereitstellung irrelevanter Informationen verringern. Dies deutet darauf hin, dass individuell angepasste Übermittlung von landwirtschaftlichen Beratungsinhalten für Kleinbäuerinnen und Kleinbauern mittels digitaler Zwei-Wege-Kommunikation praktisch umsetzbar ist. Da jedes Mal, wenn der digitale Dienst benutzt wird, zusätzliche Fragen gestellt werden können, lässt sich die individuelle Anpassung der Beratungsinhalte mittels lernender Algorithmen kontinuierlich verbessern.

Diese Dissertationsschrift trägt zum akademischen Wissen darüber bei, wie landwirtschaftliche Beratungsdienste im Globalen Süden den Herausforderungen von Ausmaß und Komplexität des kleinbäuerlichen Kontexts mit mehr methodischem Pluralismus, verstärkter Partizipation der Zielgruppe und effizienter, systematischer Nutzung digitaler Medien begegnen können. Basierend auf den vorgelegten Erkenntnissen schließt die Arbeit mit einer Erörterung weiterer Erfordernisse und Herausforderungen, die anzugehen sind, um die neuen Konzepte tatsächlich in die Praxis der landwirtschaftlichen Beratung zu integrieren. Die vermehrte Verwendung von ICT bringt allerdings nicht nur neue Kanäle zur Verbreitung von Information mit sich. Die potenzielle Diversifizierung der Beratungsinhalte sowie ein verstärkter Rückgriff auf Daten, die von Bäuerinnen und Bauern bereitgestellt werden, erfordern auch institutionellen Wandel innerhalb der Beratungsdienste. Dazu gehört etwa die Reduzierung von hierarchisch orientierten Entscheidungsprozessen durch Führungskräfte zugunsten datenbasierter, flexibler Entscheidungsfindung auf allen Ebenen, einschließlich durch das unmittelbare Beratungspersonal. Für die Zukunft ist zudem davon auszugehen, dass digitale Kommunikation zwischen der kleinbäuerlichen Zielgruppe sowie Agrarberatungsdiensten besser angenommen wird, wenn verschiedene Informations- und Kommunikationsdienste in gut gestalteten digitalen System vereint werden. Die vorliegende Dissertationsschrift macht hierzu einen Vorschlag, wie die drei hier untersuchten Konzepte in einen einzelnen digitalen Dienst für landwirtschaftliche Beratung integriert werden könnten. Auf lange Sicht können politische EntscheiderInnen erwünschte Synergien zwischen der Digitalwirtschaft und der landwirtschaftlichen Beratung erleichtern, indem zum Beispiel steuerliche Anreize oder lokale ‚Tech Hubs' als Orte der Zusammenarbeit bereitgestellt werden.

iii List of featured publications

Jonathan Steinke, Jacob van Etten, Pablo Mejía Zelan (2017). The accuracy of farmer-generated data in an agricultural citizen science methodology. *Agronomy for Sustainable Development* 37:32.

Jonathan Steinke, Majuto Gaspar Mgimiloko, Frieder Graef, James Hammond, Mark T. van Wijk, Jacob van Etten (2019). Prioritizing options for multi-objective agricultural development through the Positive Deviance approach. *PLoS ONE* 14(2):e0212926.

Jonathan Steinke, Jerusha Achieng, James Hammond, Selamawit Sileshi Kebede, Dejene Kassahun Mengistu, Majuto Gaspar Mgimiloko, Jemal Nurhisen Mohammed, Joseph Musyoka, Stefan Sieber, Jeske van de Gevel, Mark T. van Wijk, Jacob van Etten (2019). Household-specific targeting of agricultural advice via mobile phones: Feasibility of a minimum data approach for smallholder context. *Computers and Electronics in Agriculture* 162:991–1000.

iv List of Tables

5.3 Household-specific targeting of agricultural advice via mobile phones: Feasibility of a minimum data approach for smallholder context

v List of Figures

1 Introduction

5.1 The accuracy of farmer-generated data in an agricultural citizen science methodology

5.2 Prioritizing options for multi-objective agricultural development through the Positive Deviance approach

vi Acknowledgements

The work presented in this dissertation is the product of intensive cooperation and immeasurable support by a large group of people. Although put on paper by me, this dissertation took shape only thanks to a long list of persons and organizations. Sincere gratitude to you!

I wish to thank my academic and professional mentors, Jacob van Etten and Stefan Sieber, to whom I owe great part of the progress I had the chance to make over the last five years.

Many thanks to Prof. Wolfgang Bokelmann, for accepting my supervision 'last-minute', for the unconditioned support in travelling to conferences, and for offering cake and candy to the PhD students' office more than once.

Thanks to my office colleagues at Humboldt Uni, Vandreé Palacios Bucheli and Robert Cárcamo Mallen, as well as to my numerous office colleagues at ZALF for the good vibes!

Thanks to all colleagues at Bioversity for the smooth collaboration, especially to Luis Allen, Karol Araya, and Carlos Quirós.

Berta Ortiz Crespo deserves special thanks for the great collaboration on design, Ushauri, RHOMIS enumeration, and much more fun stuff in Ethiopia, Kenya and Tanzania. I have no idea what the SAIRLA project would have become without you.

Thanks to all the amazing staff at PRR (Honduras), FIPAH (Honduras), Mekelle University (Ethiopia), LWR (Kenya), and TARI-Naliendele (Tanzania) for the trust, diligence, and friendship. Above all, thank you Mainor Pavón, Pablo Mejía, Marvin Gómez, Dejene Kassahun Mengistu, Jemal Nurhisen Mohammed, Joseph Musyoka, Grace Waweru, Joy Wanza, Happy Daudi, and Gaspar Mgimiloko.

Greatest thanks to all my co-authors. The research team is stronger than the sum of its parts!

Thanks to Anna Müller (Bioversity, Turrialba) and Jeske van de Gevel (Bioversity, Nairobi), who have inspired some important thoughts for the Conclusion and Outlook sections of this dissertation, on the institutional needs for realizing digital agro-advisory.

I am grateful to Bioversity International for hosting me, as well as USAID and DFID (via the SAIRLA programme) for providing funding to the research presented in this dissertation.

Deep gratitude to my parents, Miriam and Helmut, who continue to be role models of diligence, curiosity, and cosmopolitanism throughout my life. Thank you for all the support I was fortunate to receive.

Thank you, Anne, for your support during the good and the bad times of writing a PhD thesis.

And, at last, an extended *Muchas gracias – Ahsante sana – Ameseginalehu* to all collaborating farmers in the different research sites, for the kind cooperation and key contributions to this research work.

1 Introduction

1.1 Increasing knowledge needs in smallholder agriculture

In the 21st century, smallholder farming continues to represent the main source of livelihood for the world's poor. Despite widespread growth of urban centers and the manufacturing sector, an estimated 40 percent of the developing world[1] population remain engaged in small-scale agriculture (Olinto et al. 2013). Some 479 million individual farms – 84 percent of all farms worldwide – operate on land holdings smaller than 2 ha, often at limited levels of mechanization and modern inputs, and facing persistent yield gaps (Vanlauwe et al. 2014; Lowder et al. 2016; Sheahan and Barrett 2017; Baudron et al. 2019). Notwithstanding these limitations, smallholders are believed to produce a third of global food supply on some 24 percent of global farmland (Ricciardi et al. 2018). This makes smallholder farming an important contributor to global food sufficiency.

In a rapidly changing world, however, smallholder farmers are exposed to numerous emerging pressures, as well as opportunities. Increasingly globalized value chains, changes in dietary habits, growing soil nutrient depletion and land scarcity, accelerated global migration of pests and diseases, and new environmental policy are but some of the major drivers of recent change in smallholder agriculture (Morton 2007; Hazell and Wood 2008; Pretty et al. 2010; McDonald and Stukenbrock 2016; Tariq et al. 2018). In

---

[1] None of the terms 'developing world', 'developing countries', or 'Global South' are unproblematic in their use for collectively referring to low- and middle-income countries (LMICs), which include most nations and territories of Africa, Asia, and Latin America. Though a possible alternative, 'LMIC' also reinforces a reductionist understanding of human and economic development centered on the nation-level average income level. Given the lack of fully satisfactory terminology, all three terms mentioned initially are used in this dissertation interchangeably, in awareness of their conceptual weaknesses.

addition, many regions of the developing world have been experiencing direct effects of climate change in recent years, including prolonged dry spells and increased rainfall variability, shifts in growing seasons, and increased temperatures. Recent evidence as well as projections to the future suggest that these climatic changes will cause current smallholder agro-ecosystems to suffer from decreased average yields, above all for maize (Jones and Thornton 2003; Morton 2007; Schlenker and Lobell 2010; Lobell et al. 2011; Müller et al. 2011; Challinor et al. 2014).

Under changing conditions, many smallholder farmers may sustain their livelihoods only by adapting their activities. Adaptation measures may mitigate risks, among others, through agronomy (e.g., adjusting the cropping calendar, changing the use of farm inputs, switching to better-adapted crops, cultivars, or breeds), more holistic farm re-design (e.g., the uptake of agroforestry, organic farming, 'push-pull' integrated pest management), or institutional adjustments (e.g., cooperative irrigation schemes, index insurance) (Howden et al. 2007; Pretty et al. 2011; Bryan et al. 2013; Cohn et al. 2017; Pretty 2018; Hansen et al. 2019).

These reconfigurations of farming practice, decision-making, and resource allocation, however, frequently require smallholders to acquire and apply new knowledge and skills (Pannell et al. 2006). As an example, the principle of Conservation Agriculture (CA), which requires minimum soil disturbance, has been promoted as a means to enhance small farm productivity and climate resilience in Sub-Saharan Africa (Giller et al. 2011; Corbeels et al. 2013). Successful adoption of CA means that farmers must learn to master new, complex techniques of soil management and sowing (Affholder et al. 2010), but may also cause demand for increased negotiation and new institutional arrangements around livestock feeding (Giller et al. 2009; Rufino et al. 2011).

The observed quick pace of changes affecting smallholder agriculture rules out relying on farmers' adaptation to occur by unassisted, evolutionary processes of knowledge diffusion and technological progress (Darnhofer et al. 2010). Rather, the global public goods embodied by agriculture, such as food production and landscape management, call for supporting innovation processes in smallholder farming. This may, for example, include the (co-)generation and diffusion of new knowledge, and the facilitation of individual learning processes.

1.2 Knowledge brokerage by agricultural advisory services

Agricultural extension services, also called advisory services, are in place in most countries of the developing world. Often under public governance as part of the ministry of agriculture, these organizations are mandated to provide support to farmers, for example, through information delivery and hands-on trainings, facilitation of farmer groups, or building network linkages, for example, with traders (Davis 2008; Benson and Jafry 2013; Leeuwis 2013). Private extension services offered by development NGOs or agricultural input companies frequently operate alongside public providers (Davis 2008).

Advisory staff typically assist farmers in problem-solving through field visits, promote and recommend agricultural technology[2], and carry out training courses at extension facilities. The overall approach to extension delivery, however, varies greatly between countries and regions: The widespread training-and-visit model (Feder et al. 1986), for example, focuses on 'contact farmers' (also, *model* or *lead* farmers), counting on further farmer-to-farmer knowledge transfer (Niu and Ragasa 2018; Taylor and Bhasme 2018). In Farmer Field Schools (van de Fliert et al. 1995), advisors regularly meet groups of farmers for problem-based, hands-on trainings about agricultural topics (Friis-Hansen and Duveskog 2012).

While these efforts have arguably contributed to increased farm productivity and sustained incomes at many places (Davis et al. 2012; Wossen et al. 2017), in some cases, technology adoption and improvements in well-being have been disappointingly weak or absent (Faure et al. 2012; Rejesus et al. 2012). From the 1980s on, such observed weaknesses of agricultural extension have stimulated criticism of the linear paradigm of technology transfer from research to farmers, mediated by advisory services. Researchers and policymakers have hence increasingly emphasized the direct involvement of farmers in the generation, specification, and dissemination of knowledge and innovation (Farrington and Martin 1988; Sumberg et al. 2003; Hoffmann et al. 2007; Knickel et al. 2009; Lubell et al. 2014).

[2] In the context of agriculture, 'technology' refers to both *"[t]he application of scientific knowledge for practical purposes, especially in industry"* and *"[m]achinery and equipment developed from the application of scientific knowledge"* (alternative definitions given by Oxford Dictionaries, online at *en.oxforddictionaries.com*). Agricultural technology may therefore mean knowledge, such as good agricultural practice, as well as embodied technology, such as improved seeds.

Modern extension methodologies have thus integrated the facilitation of context-bound, farmer-participatory agricultural research to identify suitable solutions for local problems, for example, through 'Local Agricultural Research Committees' (CIALs) in Latin America (Ashby et al. 2000; Neef and Neubert 2011; Humphries et al. 2015). More recently, there has been increased recognition of the importance of engaging the perspectives of other food system actors, beyond researchers and farmers, in innovation processes in agriculture (Kristjanson et al. 2009; Spielman et al. 2009; Kilelu et al. 2011). Most notably in Africa, a more systemic approach to agricultural extension has hence been explored. Facilitated multi-stakeholder 'innovation platforms' employ dialogue, negotiation and joint reflection to support the co-creation of social, technical, or economic innovation among, for example, farmers, researchers, policy-makers, NGOs, policy-makers, and traders (Kilelu et al. 2013; Lamers et al. 2017; Schut et al. 2018; Schut et al. 2019).

Over the last decade, however, existing limitations of these new methodologies have also become apparent. Providing useful agricultural advisory to a large and heterogeneous farming population requires approaches that can efficiently be implemented at scale, while addressing the diverse conditions at different farms. In practice, boundary-spanning innovation platforms have rarely been scaled beyond local pilot projects, and have experienced only limited integration into national extension institutions (Schut et al. 2016). Recent research progress is already demonstrating how extension services can harness to power of large-scale farmer participation and intensive interaction without the need for resource-intensive group facilitation. Modern information and communication technology (ICT), offering new opportunities for rapid, systematic communication, is key to these new approaches (Munthali et al. 2018; Nelson et al. 2019).

1.3 Research objectives

The overall aim of this dissertation is to contribute to methodology development for agricultural extension in developing country smallholder context. The analysis combines theory from two emerging fields of innovation – large-scale farmer participation and ICT – to guide research towards three interrelated objectives:

- To describe novel approaches for addressing common deficiencies of agricultural extension in the developing world
- To test the practical viability of these approaches through proof-of-concept studies

- To deliver empirical insights on the usefulness of these approaches for improving agricultural extension services

1.4 Synopsis of this work

This cumulative dissertation addresses its research objectives by a three-pronged research approach: For three selected challenges of agricultural extension, three potential solutions are suggested. Then, three peer-reviewed publications provide empirical evidence on the feasibility and usefulness of these concepts to address the challenges. Figure 1 visualizes this research concept, including the theory inputs that led to the choice of solution concepts.

In the following document, Chapter 2 includes the research framework that guided research design and analysis. First, the chapter outlines three key methodological shortcomings that tend to challenge the performance of agricultural extension services in the developing world (Section 2.1). Then, two emerging opportunities for the development of agricultural advisory services are presented: Large-scale farmer participation (Section 2.2) and the use of modern ICT (Section 2.3). Informed by synergies between these two broad opportunities, three concrete novel solution concepts are described (Section 2.4). Each solution concept is suggested as a methodological innovation to address one of the identified shortcomings of agricultural extension.

Chapter 3 lists three research questions. Each question relates to the potential of one specific solution concept to address the respective shortcoming of agricultural extension.

Chapter 4 provides a brief overview of research locations, activities, and methods applied.

Chapter 5 presents results. For each research question, one peer-reviewed publication presents empirical evidence on the potential of the suggested solution concept, and discusses respective limitations.

Chapter 6 synthesizes the results, acknowledges research limitations and remaining open questions, and draws conclusions for future methodological development of agricultural advisory services in the developing world.

Chapter 7 concludes the dissertation by providing an optimistic outlook on expected future development in the field.

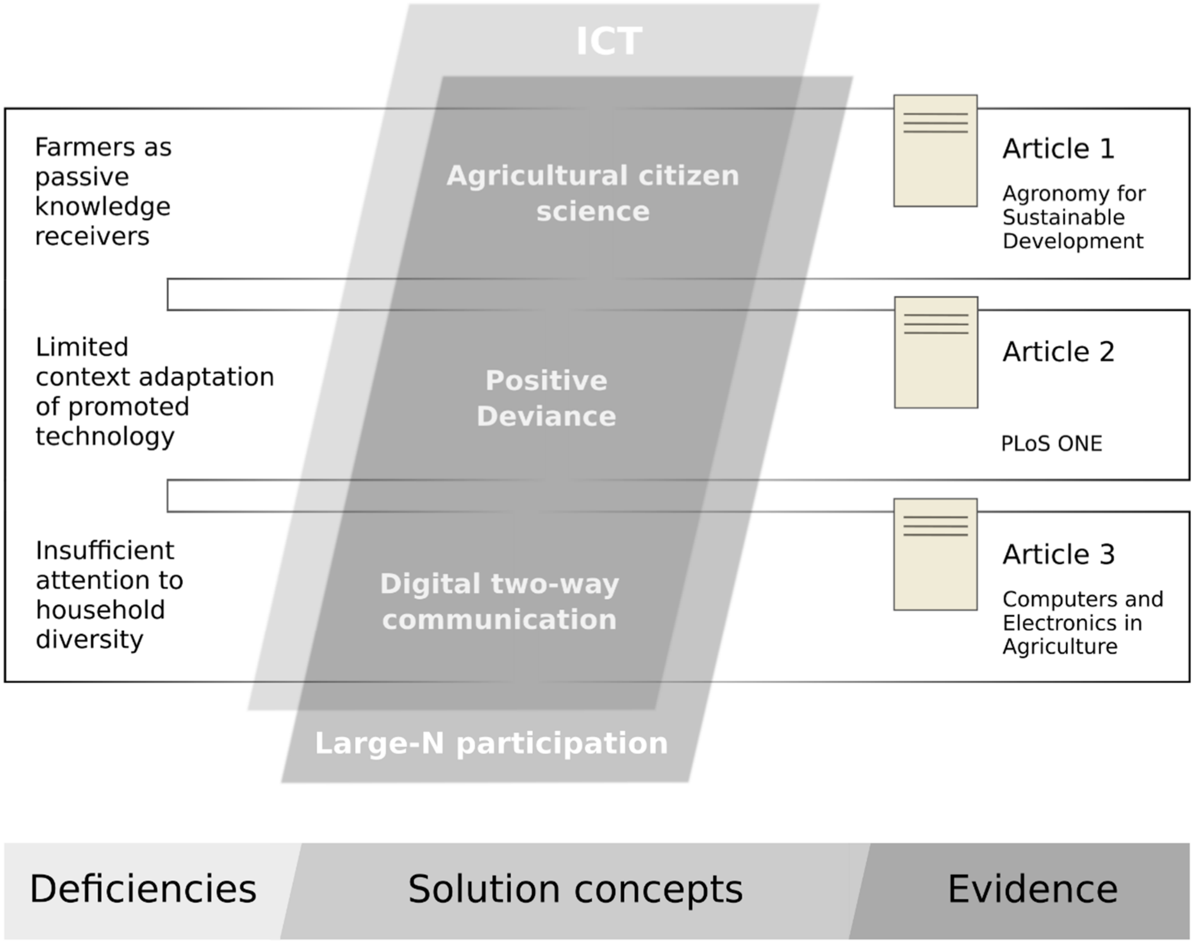

Figure 1: Research concept: Selected deficiencies of agricultural extension, proposed solution concepts based on synergies between large-scale participation and ICT use, and empirical evidence on their feasibility

2 Research framework

2.1 Three deficiencies of agricultural extension

Agricultural advisory services worldwide are characterized by strong diversity. Providing agencies differ in governance and funding (e.g., public, fee-based, cooperative), in methodologies, as well as in their focus, such as integrated farming systems vs. single commodity value chains. Existing diversity in farming systems and rural livelihoods across and within developing countries implies there is no gold standard for agricultural extension services. As every approach to agricultural extension has strengths and weaknesses, depending on context, 'pluralism' in advisory services is a necessity (Birner et al. 2009). Nonetheless, critical appraisal of the existing experiences highlights several shortcomings that are widespread within current agricultural advisory services. Despite often locally adapted, diverse extension agendas and methods, multiple systemic challenges frequently affect the outcomes of advisory services across contexts (Anderson and Feder 2004; Davis 2008; Faure et al. 2012).

A pivotal challenge limiting the effectiveness of extant extension methodologies is the *Scale and Complexity* of smallholder farming context. Anderson and Feder (2004:45) outline this problem:

> "Thus, the number of clients who need to be covered by extension services is large, and the cost of reaching them is high. Adding to the challenge, farmers' information needs vary even within a given geographical area because of variations in soil, elevation, microclimate, and farmers' means and capabilities. The large size of the clientele means that only a small number of farmers can interact directly with extension agents."

Public and private extension services in the developing world can face a variety of concurrent challenges, including lack of political support and unreliable funding. But ongoing developments, such as population growth, increased globalization of commodity markets, and climatic changes, play in favor of continued increases in the total *Scale and Complexity* of smallholder farming context. On the whole, these changes can gradually aggravate the difficulty of delivering adequate agricultural advisory to smallholder farmers via existing methodologies. While many observed limitations of agricultural advisory services can be addressed through responses in policy, funding, and internal governance, the particular challenges associated with *Scale and Complexity* relate closely to the methodologies directly employed by extension services. Therefore, interest in exploiting recent technological development to advance extension methodologies is high.

To ensure close problem-orientation, this work focusses on three major aspects that frequently challenge agricultural advisory delivery in the developing world: First, the high numbers of targeted farmers imply that extension services tend to engage farmers merely as passive receivers of existing knowledge, rather than encouraging the active generation and contribution of knowledge. Second, strong heterogeneity of smallholder farming systems and socio-economic contexts causes extension services to focus on few 'best-bet' technologies. Thus, agricultural knowledge and technology is often promoted based on stationary trials or success stories at other locations, overriding the need for identifying solutions that are viable in local smallholder farming context. Third, even when extension services hold knowledge about appropriate solutions for diverse contexts, the high number of farms, combined with their strong heterogeneity, often precludes adequate disaggregation and household-specific targeting of agricultural advice. These shortcomings of extension services in view of increasing *Scale and Complexity* are described in more detail below.

First deficiency: Farmers as passive knowledge receivers

A core mandate of most agricultural extension services worldwide is to provide farmers with knowledge and information that is useful and applicable to their operations and decision-making. This can include, for example, the supply of agronomic advice or market price information. Access to technological knowledge and relevant agro-information may often be an important driver of adaptation and development in smallholder farming. But the practical value of advice can be limited by the high degrees of uncertainty and unpredictability that characterize smallholder agriculture. As an example, extension services may recommend standard fertilizer dosages. But changes in input and output prices, climate, or crop variety choice may subsequently affect the applicability of those recommendations. Coping with uncertainty and responding to rapidly changing pressures and opportunities may require practical skills, heuristics, tacit forms of knowledge, and the ability to improvise, rather than 'having' information and static knowledge (Suppe 1987;

Gigerenzer and Gaissmaier 2011; Šūmane et al. 2018). Making diverse experiences, for example through experimentation, can strengthen farmers' abilities to make accurate observations of new situations, and to adapt by drawing the right inferences from experience (Richards 1989; Richards 1993; Darnhofer et al. 2010). Technological change on small farms is also more likely to happen when farmers participate in technology development (Johnson et al. 2003, Sumberg et al. 2003). As pointed out by Glover et al. (2016:4), *"technology is something people do, make or remake, not something they receive or adopt."*

Extension formats that emphasize farmer experimentation and the creation of experiential knowledge have been in use since at least the 1980s. Farmer Field Schools, for example, supply local farmer groups with agricultural technology, such as improved seeds, and encourage group-based experimentation and experiential learning (van de Fliert et al. 1995; Davis et al. 2012; Friis-Hansen and Duveskog 2012). With a stronger focus on knowledge generation and technology development, 'Local Agricultural Research Committees' (CIALs for the Spanish acronym) engage in designing, implementing, and analyzing systematic agricultural experiments (Ashby et al. 2000). In the past, CIAL research, facilitated by private extension agencies, has led to the release of farmer-bred crop varieties which outperformed commercial varieties under farmer management. But in addition, participants acquired skills and were empowered for creative, independent problem-solving (Classen et al. 2008; Humphries et al. 2015).

In practice, however, the resource-intensity of these group-based approaches means they are rarely used at scale: Farmer-participatory experimentation and joint, experiential learning processes often rely on costly, continuous facilitation of organized farmer groups. Therefore, these formats are usually restricted to small shares of the farming population (Ashby et al. 2000; Neef and Neubert 2011). Most smallholder farmers in the developing world continue to interact with extension services, if at all, by one-way processes of information delivery (Davis 2008; Faure et al. 2012). To address the need of the wider farming population for building adaptive capacity through exposure to diverse experiences, there is a need for scalable, inclusive methodologies that facilitate farmer experimentation and experiential learning at scale, without presupposing group membership or existing research capacities.

Second deficiency: Limited context adaptation of promoted technology

Agricultural extension services are typically expected to increase smallholder productivity and resilience, and, more generally, the wellbeing of rural households. To achieve these goals, specific agricultural technologies are promoted to farmers, such as farm inputs and agronomic or post-harvest practices. Extension services that collaborate closely with the local research sector often emphasize recent research products, such as newly released crop varieties (Asfaw et al. 2012; Verkaart et al. 2017). In contrast, many private extension services, often receiving international funding, emphasize the introduction and out-scaling of 'best-bet' technologies that have achieved desired results at other places with similar conditions (Muthoni et al. 2017; Notenbaert et al. 2017).

Despite their promise, the introduction and promotion of new technologies has not always generated the expected positive outcomes (Feder et al. 1985; Moser and Barrett 2003; Arslan et al. 2014). Low adoption rates or high levels of eventual dis-adoption suggest that some features of the technology are incompatible with dominant farming systems or local culture, or unfold unexpected disbenefits under smallholder field conditions (Pannell et al. 2006; Snyder and Cullen 2014; Grabowski et al. 2016). As an example, the System of Rice Intensification (SRI) has the potential to increase yields at low level of external inputs. In Madagascar, however, SRI was found to be associated with increased labor burden at peak-labor periods of the agricultural season, thus inhibiting widespread uptake of the technology (Moser and Barrett 2003). Similar limitations have been found to restrict adoption of soil-fertility increasing Conservation Agriculture (Corbeels et al. 2013; Andersson and D'Souza 2014; Grabowski et al. 2016).

Incompatibilities with target context are hard to avoid where technologies are not being developed under the specific constraints and requirements of local farmer management. This challenge affects both stationary research products and technologies 'imported' as promising solutions from elsewhere. To come up with locally suitable solutions, farmer-participatory or farmer-led agricultural technology development has become widespread in recent decades (Farrington and Martin 1988; Sperling et al. 2001; Joshi et al. 2012). But participatory research approaches also have limitations: Group-based methodologies often use a low number of shared research locations, which do not always reflect the conditions and constraints at individual farms. Differences in biophysical conditions and labor availability between shared research plots and individual farms sometimes limit the potential for sustained adoption of new technologies (Misiko 2013). The limitations of existing methods used by extension services for selecting and prioritizing promoted technologies suggest that better methods are needed for determining which agricultural development options are likely to be successful and acceptable under local smallholder farm conditions.

Third deficiency: Insufficient attention to household diversity

Smallholder farming systems in the Global South show strong heterogeneity at the regional, community, household, and individual plot levels. Important dimensions of diversity relate to, e.g., household productive assets, such as land and livestock (Lowder et al. 2016; Berre et al. 2019) and access to labor (Kuivanen et al. 2016b), soil quality (Tittonell et al. 2007b; Tittonell et al. 2010), market opportunities (Barrett 2008), cropping systems (Frelat et al. 2016; Ritzema et al. 2017; Wichern et al. 2018), agro-climatic conditions (van Wart et al. 2013), and farmers' aspirations, experiences, and attitudes (Dorward et al. 2009; Mausch et al. 2018; Verkaart et al. 2018). This diversity implies large differences in individual adoption potential and potential benefits of specific agricultural technologies, and rules out the use of blanket recommendations by extension services. Rather, in recent years, efforts have been made to deliver agricultural advice that is tailored to farmers' contexts, including their specific opportunities, constraints, and information needs. For example, integrated farming system models can be used to simulate the effects of different options of technological change at the household level. These models, however, hardly fit in with the daily routines of extension services working with smallholder farmers in the Global South, as they require large sets of household data, considerable time dedicated to individual farms, and advanced academic skills (Bernet et al. 2001; Woodward et al. 2008; Vayssières et al. 2011).

In practice, extension services often use heuristic household categorizations to target individual farmers with different information and advice. Simple typologies distinguish, for example, between growers and non-growers of certain (cash) crops. More detailed farm typologies are widespread, often assigning households to one of multiple (usually 4 to 7) stereotypical production systems, such as 'livestock-oriented farms', 'crop-oriented farms', or 'part-time farmers' (Tittonell et al. 2010; Kuivanen et al. 2016a; Lopez-Ridaura et al. 2018; Berre et al. 2019). The concept of 'socio-ecological niches' recognizes the importance of further internal heterogeneity within household types, and calls for participatory, iterative specification of typology-based 'baskets of technology options' (Ojiem et al. 2006; Descheemaeker et al. 2019). Farmer-participatory methods of technology prioritization that allow tailoring advice to context include, for example, rapid appraisals and ex-ante impact assessments (Schindler et al. 2016; Mwongera et al. 2017).

Despite the availability of different methods to inform tailored, diversified advisory and their potential to increase technology adoption, household-specific targeting of advice by agricultural extension service is still the exception rather than the rule (Faure et al. 2012). In many cases, necessary targeting exercises are relatively costly for resource-restricted extension agencies. Developing detailed statistical farm typologies and matching diverse options-by-context requires large sets of reliable household data as well as advanced academic skills. More participatory, qualitative methods require substantial time commitments, and results are often hardly scalable beyond the participating farming community.

These limitations indicate a trade-off between the cost of employing specific targeting methods, and the need for responding to the diverse knowledge and information needs of a heterogeneous farming population. More effective methods for tailoring agricultural advisory may help to provide more adequately disaggregated agricultural advice with given resources.

2.2 Opportunities of large-scale farmer participation in agricultural research and extension

Over the last decades, public participation has become a ubiquitous component of agricultural research and development (Neef and Neubert 2011). The experience that stationary research and linear transfer-of-technology tended to lead to limited adoption and frequent dis-adoption of technologies has motivated the development of various methodologies for involving farmers actively in research. These approaches harness farmers' diverse perspectives, experiences, and creativity to foster the development of technology and build knowledge that is adapted to the constraints, needs, and preferences of local smallholder farmers. Well-established examples include the co-development of new, farmer-preferred crop varieties or locally acceptable innovations in farming systems design and agronomy (Farrington and Martin 1988; Sumberg et al. 2003; Bezner Kerr et al. 2007; Hoffmann et al. 2007; Ceccarelli et al. 2009; Méndez et al. 2017; Tariq et al. 2018).

The boundaries to agricultural extension, both in terms of goals and methodologies, are often fluid. For example, extension services frequently engage 'model farmers' or 'demonstration farms' under smallholder management in order to adapt technologies to local context and to facilitate horizontal diffusion through increased observability (Niu and Ragasa 2018; Taylor and Bhasme 2018). Such farmer-to-farmer extension formats benefit from high local adaptation of knowledge and the inclusion of tacit knowledge, but can also suffer from biases and information loss (Niu and Ragasa 2018). Group-based experimentation formats, such as Farmer Field Schools (van de Fliert et al. 1995) or participatory variety selection (Witcombe et al. 1996; Witcombe et al. 2005), additionally emphasize interactions between farmers as well as experiential learning, often facilitated by extension personnel. Such group-based participatory extension methodologies have generated successes in terms of learning, technology diffusion, and knowledge generation (Johnson et al. 2003; Knook et al. 2018). But there are also inherent limitations: Group deliberations, for example, can marginalize women (Cornwall 2003), risk overriding minority opinions (Brodbeck et al. 2002), and results may be biased by the identity of the facilitator

(Humphreys et al. 2006). Moreover, because these approaches often rely on joint observation and experimentation on collective plots or demonstration farms, results are not always easily applicable to the individual farms of participants, which may be subject to different biophysical conditions or labor constraints (Misiko 2013). Lastly, the focus on organized, facilitated groups and specific locations implies that some farmers – for example, those residing in most remote locations or with highest workloads, often women – are less likely to participate (Meinzen-Dick and Zwarteveen 1998; Feder et al. 2010).

Recent technological development has created new opportunities for overcoming some of the limitations of group-based extension format by engaging potentially large numbers of farmers on an individual base through modern ICT (Munthali et al. 2018). In the environmental sciences, methodologies for facilitating large-scale individual participation in collective research processes, as well as digital platforms for horizontal sharing of experiences and knowledge, have been established over the last decade (Dickinson et al. 2012; Newman et al. 2012). The emerging application of such technology-mediated large-scale participation methodologies to the agricultural field combines the advantages of facilitated individual on-farm experimentation (e.g., local learning under real-life farm conditions, no transaction or travel costs for group research) with the strengths of group-based methodologies (e.g., larger numbers of technologies tested, replication of trials, social learning through farmer deliberations) (van Etten 2011; van Etten et al. 2019b).

Major challenges for the design of large-scale farmer participation relate to the efficient coordination of data and knowledge flows within the knowledge system. Moreover, incentivizing autonomous participation can be challenging in absence of direct interaction with researchers and extension agents (Ortiz et al. 2011; Beza et al. 2017). This suggest that large-scale participation approaches need to be designed in such way that (i) large numbers of farmers can contribute useful data with minimum training, and (ii) advisory feedback implies added value for participants, exploiting the power of large (farmer-generated) datasets and providing meaningful insights beyond the individual farm level (Nelson et al. 2019). Knowledge systems that connect many individual participants may benefit strongly from the emerging opportunities of digital platforms and affordable end-user devices for two-way communication, such as mobile phones.

2.3 Opportunities of modern ICT for agricultural extension services

Recent technological and infrastructural development across many countries of the Global South has created new opportunities for systematic communication around smallholder agriculture. It is expected that some of the existing limitations of agricultural extension (see Section 2.1 above) can be addressed through improved information flows between the different stakeholders of a complex knowledge system, including farmers, researchers, and extension agencies (Duncombe 2015; Deichmann et al. 2016). In particular, the continued growth of mobile phone availability and network connectivity in most developing countries may enable new, efficient forms of extension delivery to an often geographically dispersed farming population (Aker and Mbiti 2010; Aker 2011; Nakasone et al. 2014). In the African continent, for example, there were an estimated 76 active mobile phone subscriptions per 100 inhabitants in 2017 (ITU 2018). 44% of the population in Sub-Saharan Africa had a subscription of their own (GSMA 2018). Consequently, public agricultural extension agencies, NGOs, as well as private for-profit companies worldwide have created many mobile phone-mediated information services for smallholder agriculture over the last decade. Examples include SMS-based market information systems, call centers for agronomic advice, or automated weather forecast hotlines (Aker et al. 2016; Baumüller 2018).

Because delivering information via ICT is cheaper than through visits by advisory staff, extension services can realize higher numbers of extension contacts to farmers, using given budgets (Aker 2011). First-generation digital agro-information services have, therefore, often focused on large-scale one-way information dissemination in three areas: Market information (e.g., day-to-day market prices), agro-climatic information (e.g., short-term weather forecasts, seasonal climate forecasts), and agronomic farming advice, such as variety recommendations (Aker et al. 2016; Baumüller 2018). But mobile phones have greater functionality than radios: Beyond information reception, modern ICT offer many opportunities to collect, organize, and utilize data and knowledge inputs from farmer-users (Minet et al. 2017; Munthali et al. 2018). Recent initiatives have started to leverage the potential of ICT for enabling new forms of knowledge (co-)generation, information disaggregation and tailoring, and horizontal knowledge sharing between farmers (Patel et al. 2010; Eitzinger et al. 2019).

For example, employing digital two-way communication between farmers and agricultural extension allows delivering tailored advice in response to farmers' individual data inputs. In the *AgroDecisor* smartphone application, soybean producers answer ten questions, for example, about their farming practice or recent rainfall events. Through a simple algorithmic scoring system, the service then returns immediate plot-specific recommendations on pesticide application (Carmona et al. 2018). But modern ICT can also facilitate more complex network interactions between farmers and advisors: Another example for the novel ways in which modern ICT can support agricultural extension is *GeoFarmer*, a digital service combining a smartphone application with a more basic technology, interactive voice response (IVR). This environment allows farmers and agricultural advisors to submit georeferenced farm observations, to ask, discuss, and mutually answer diverse questions, and to track changes in on-farm decision-making (Eitzinger et al. 2019).

Investments into the development of new digital agro-advisory services can be justified by increased efficiency over conventional extension formats (e.g., through more extension contacts per farmer) or higher levels of inclusion (e.g., by allowing remote or marginalized farmers to access advice that used to be a privilege of wealthier farmers) (Deichmann et al. 2016). Both criteria, efficiency and inclusion, imply that ICT-based solutions should aim at involving large number of farmers. But in addition, 'direct network effects' cause some digital services to unfold increasing benefit with increasing numbers of users (see Gawer 2014). In particular, larger diversity among the farmers contributing data or knowledge inputs to a service may often improve the overall output, for example, allowing more finetuned targeting of advisory messages, or leading to a more comprehensive body of questions and answers.

2.4 Novel solution concepts

Agricultural extension in the Global South has great potential to benefit from synergies between large-scale farmer participation and the use of modern ICT. This dissertation explores three new methodological concepts that aim at addressing specific deficiencies of agricultural extension by combining the opportunities of large-scale farmer participation and ICT.

Agricultural citizen science

'Citizen science' refers to a diversity of systematic scientific processes in which non-professional scientists, often volunteers, become involved at any stage. In the environmental and biological sciences, the great potential of engaging nature enthusiasts and science aficionados for data collection and analysis has been harnessed for more than a decade, already (Cooper et al. 2010; Hand 2010; Conrad and Hilchey 2011; Dickinson et al. 2012). Successful citizen science projects build on the concept of 'crowdsourcing', where participants individually make small contributions, and a central management system carries out (meta-)analysis, feeding back collective results to the individual contributors. As an example, volunteer bird watchers in the US regularly submit georeferenced bird observations to the *eBird* online portal, jointly generating powerful and rapidly available data for ornithological research (Sullivan et al. 2009; Sullivan et al. 2014). In the *Foldit* online game, players solve puzzles about protein structures and collectively generate results that have accelerated bioengineering research (Khatib et al. 2011; Eiben et al. 2012).

More recently, systematic crowdsourcing research methodologies have also gained momentum in the agricultural sciences (Minet et al. 2017; Ryan et al. 2018). By dividing agricultural experiments into many small, farmer-run 'micro-trials', research activities can be carried out under full farmer management and in diverse environmental conditions, speeding up the generation of research output (van Etten et al. 2019b; Fadda and Van Etten 2019). Especially for plant breeding, generating and compiling data from smallholder on-farm trials is promising, as realistic representations of farmers' diverse preferences and on-farm conditions, including constraints, are crucial for sustained adoption of new crop varieties (Misiko 2013; Almekinders et al. 2019).

The growing availability of mobile phones makes it possible to collect observations from farmers without close researcher supervision or on-site interaction (Dillon 2012; Daum et al. 2018). At the same time, involving large numbers of individual contributors in research through crowdsourcing can imply a trade-off between data quality and quantity (Kremen et al. 2011; Hochachka et al. 2012; Hunter et al. 2013). Therefore, agricultural citizen science methodologies require data collection and analysis procedures that lead to meaningful scientific results despite limited opportunities for data quality control.

Positive Deviance

Agricultural development organizations and extension services across the developing world have frequently been confronted with low adoption rates and widespread post-project dis-adoption of innovations targeted at resource-poor farmers (Moser and Barrett 2003; Arslan et al. 2014; Wanvoeke et al. 2015; Grabowski et al. 2016). Often, disappointing adoption dynamics can be explained by mismatches between presumed and real use context, for example, when practices show different benefits and disbenefits under individual smallholder management than under experimental management. Moreover, the heterogeneity of smallholder context and local culture can hinder wide adoption of promising agriculture technology (Henrich 2001; Crane 2010).

The 'Positive Deviance' approach is an alternative strategy for development agencies to select interventions. It refers to *"the observation that in most settings a few at risk individuals follow uncommon, beneficial practices and consequently experience better outcomes than their neighbours who share similar risks."* (Marsh et al. 2004:1177). Positive Deviance suggests that certain households achieve stronger livelihood performance than other households with similar resources and challenges. These households are likely to use available resources more efficiently and are potential sources of information about locally viable innovative practice. The approach was originally developed by nutritionists in the 1990s: First, using a quantitative survey, they identified poor households with remarkably healthy children. Then, they re-visited those households to identify any 'deviant' hygiene and child feeding practices that might explain their relative success. Subsequently, these locally suitable practices were promoted to other households in similar resource and cultural context (Sternin et al. 1998; Mackintosh et al. 2002).

Learning from surprisingly successful, innovative farmers has great potential for agricultural development initiatives (Biggs 2008; Pant and Hambly Odame 2009; Modernel et al. 2018). Given the strong diversity of smallholder context, identifying different options for different socio-ecological niches is imperative (Descheemaeker et al. 2019). Diverse positive deviants can demonstrate 'what works' in very specific context and can thus inform the prioritization of adaptation measures for similar context. But to identify useful options for heterogeneous farming households, the Positive Deviance approach needs to explore the performance and practices of equally diverse farming households. By definition, positive deviants – the relatively best performers – are rare. Therefore, a larger pool of farmers who contribute data and knowledge may allow analysis that generates more finetuned options-by-context.

The new communication opportunities offered by modern ICT could facilitate the use of the Positive Deviance approach by extension agencies in various ways: For example, remote survey data collection may continuously build and update the database of household performance that is needed to identify positive deviants (Hartung et al. 2010; Dillon 2012; Hoogeveen et al. 2014). For out-scaling innovations, digital media and platforms can support knowledge sharing between positive deviants and suitable target households at other locations, for example, through voice messages or instructive videos (Gandhi et al. 2009; Patel et al. 2013). To date, systematic application of the Positive Deviance approach for informing agricultural development efforts has not been explored empirically.

Digital two-way communication

Agricultural extension services in the developing world use a large diversity of media for communicating contents, including leaflets, posters, radio, or TV broadcasts (Niu and

Ragasa 2018; Bentley et al. 2019). These conventional mass media deliver agricultural advisory messages through linear communication from extension agencies to farmers. Modern ICT, however, also allow systematic feedback and dialogue, blurring the lines between who 'sends' and who 'receives' information. The proliferation of digital end-user devices, i.e. mobile phones, even among resource-poor farmers opens opportunities for harnessing two-way communication for agricultural development efforts.

In the context of agricultural extension, digital two-way communication between farmers and (partly) automatized advisory applications[3] has a variety of possible uses. By empowering farmers to provide and share information inputs, request specific information, or select preferred communication partners, two-way communication can lead to a more successful advisory experience (Jones and Kondylis 2018). For example, the *M-Farm* crop trading platform allows Kenyan farmers to make offers and accept bids from buyers through structured interactions by SMS (Baumüller 2015). In India, a question-and-answer forum based on short voice messages, *Avaaj Otalo*, lets farmers and advisors use their phones to engage in horizontal, interactive knowledge exchange and sourcing of peers' experiences (Patel et al. 2010). And in Argentina, Soybean farmers may use the aforementioned *AgroDecisor* app for decision support on pesticide application (Carmona et al. 2018). This kind of tailored advice is likely to be more useful than broad guidelines provided through linear communication, without feedback mechanisms.

The use of digital two-way communication for targeting agricultural advice is receiving increased attention in recent years. Mobile phone-enabled services allow collecting household-, farm-, or plot-level data from farmers, in order to return correspondingly selected or customized contents. It has been suggested, for example, to crowdsource data on crop management and plot-level outcomes (such as yield) from many farmers, allowing the feedback of highly granular advice based on local evidence (Cock et al. 2011). But to be useful in practice, data collection from farmers must be rapid and avoid lengthy, tedious enumeration (Rosenstock et al. 2017). Various solutions to the trade-off between rapid data collection and the usefulness of tailored advice exist, but there is no consensus yet on their viable implementation in two-way agro-advisory applications.

[3] Contrary to everyday speech, the term *application* in this dissertation does not refer to smartphone-based software exclusively. Rather, an advisory application can be any service that can be operated by farmers and that supports any type of information exchange through any technological interface. This may include smartphone applications, but also automated call-in hotlines, SMS subscription services, and more.

3 Research questions

Based on three identified deficiencies of agricultural extension services (Section 2.1), this dissertation proposes three solution concepts (Section 2.4). To explore the respective potentials of these concepts for addressing the shortcomings of extension and to generate evidence on their practical viability, three research questions will be answered.

Research question 1:

> How can ICT-mediated agricultural citizen science help to involve large numbers of smallholder farmers in knowledge generation?

To adopt agricultural citizen science as a tool in collaborative agricultural research, clarity is needed on the possibility of generating valid and useful research evidence based on farmer-generated data. Therefore, the research in this dissertation focuses on the accuracy of data in an agricultural citizen science methodology. Information about accuracy will allow making statements about viable scenarios under which citizen science can be employed for meaningful and robust agricultural research.

Research question 2:

> How can the Positive Deviance approach help to identify locally suitable innovation using ICT-mediated data inputs from many smallholder farmers?

The systematic application of research methods that rely on Positive Deviance is a novelty in agricultural development. This dissertation develops and tests a step-wise procedure, adapting the Positive Deviance approach for agricultural development. Conclusions are drawn on the potential for future use of the method by agricultural advisory services.

Research question 3:

> How can digital two-way communication be employed to target agro-advisory messages to heterogeneous smallholder farmers?

Given the rapid growth of mobile telephony in the developing world, this dissertation focuses on opportunities associated with the use of farmers' individual mobile phones. To generate actionable insights for future endeavors of targeting agricultural advisory messages through the mobile phone, a novel procedure for simple two-way communication is described. Through an empirical test, practical implications and value-addition of this targeting approach over linear knowledge transfer are explored.

4 Research activities

This dissertation presents research evidence that was generated through multiple interconnected steps of fieldwork and desk analysis. Activities are described in detail in the methods parts of Sections 5.1 – 5.3 (below). For an overview, the major steps in the research process, with emphasis on data collection, are outlined here briefly. All research activities were undertaken as part of projects supervised by Dr. Jacob van Etten (Bioversity International). Funding for research in Honduras was provided by USAID to Bioversity International through grant AID-OAA-F-14-0035 'Crowdsourcing crop improvement: Evidence base and outscaling model'. Research in Eastern Africa (Ethiopia, Kenya, Tanzania) was carried out within the 'What works where for which farmer' project. Funding was provided to Bioversity International by UK Aid from the UK government through the Sustainable Agricultural Intensification Research and Learning Alliance programme (SAIRLA).

Crop variety ranking experiments (Honduras)

In various parts of Honduras, Bioversity International has been collaborating with two local NGOs, PRR[4] and FIPAH[5], in the implementation of on-farm variety selection trials for common bean (*Phaseolus vulgaris* L.). These activities explored the possibilities and practical challenges of executing large-N farmer-participatory variety selection through *tricot*, an agricultural citizen science methodology. To generate insights about the accuracy of farmers' observations, 35 farmers (women and men) at five sites in four regions (Figure 2) were asked to perform tricot-style observations. In this experiment, each participant individually observed a small plot with three different varieties in vegetative state, and then ranked the varieties according to multiple selection criteria. The same observations were also collected from a trained agronomist, to determine the ranking that would be accepted as correct. The mean recorded deviations between the farmers'

[4] *Programa de Reconstrucción Rural*

[5] *Fundación para la Investigación Participativa con Agricultores de Honduras*

and the agronomist's observations were then used to assess the accuracy of farmers' observations

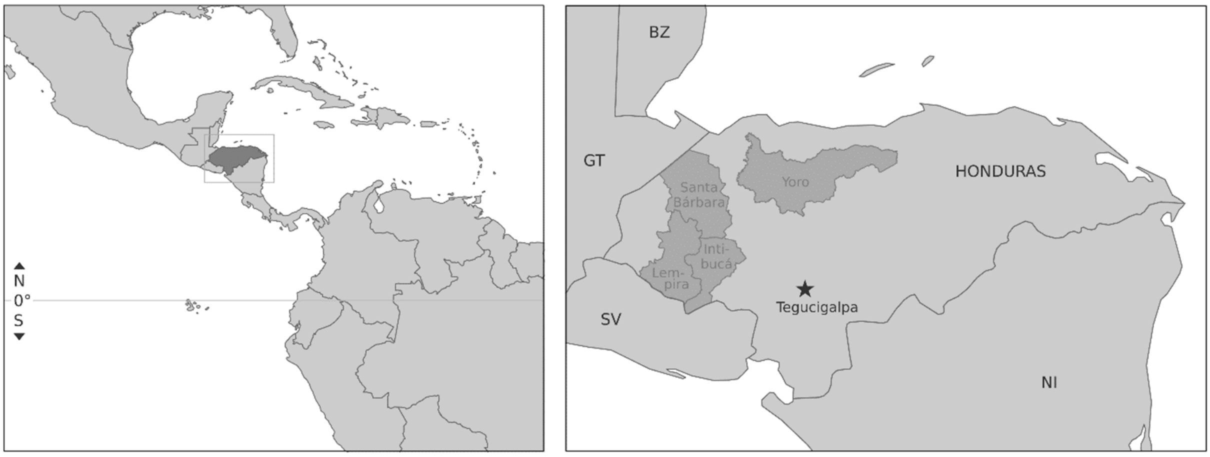

Figure 2: Central America and the Caribbean region with Honduras in dark grey (left). Location of the four research regions in Honduras, where ranking experiments were carried out (right). Neighboring countries are marked with ISO two-letter codes. Spatial data retrieved from *gadm.org*.

Household surveys (Ethiopia, Kenya, Tanzania)

As part of the 'What works where for which farmer' project led by Bioversity International, a standardized quantitative survey was administered to the heads of selected smallholder farming households in Ethiopia (Tigray region), Kenya (Makueni County), and Tanzania (Southern Agricultural Zone) (Figure 3). These surveys were variations of the *Rural Household Multiple Indicator Survey* (RHOMIS) format (Hammond et al. 2017). Enumeration efforts were led by local project partners Mekelle University (Ethiopia), Lutheran World Relief (Kenya), and TARI-Naliendele (Tanzania). The surveys provided a multifaceted characterization of individual households, for example, by their resource endowments, farming system, and agricultural and dietary performance. These household characterizations were crucial inputs to the research on both Positive Deviance and digital two-way communication. In total, 249 households were successfully surveyed in Ethiopia, 316 in Kenya, and 521 in Tanzania.

Data-driven identification of positive deviants (Ethiopia, Kenya, Tanzania)

The Positive Deviance approach has most commonly been used to identify favorable behavior regarding hygiene, health, and nutrition. One goal of the research presented here was to adapt the Positive Deviance approach to multi-objective agricultural development. To this end, positive deviants were defined as having Pareto-optimal relative performance regarding five key dimensions of agricultural development. To identify these positive

Figure 3: Southern and Eastern Africa with research focus countries in dark grey (left). Location of research regions in Ethiopia, Kenya, and Tanzania (right). Neighboring countries are marked with ISO two-letter codes. Spatial data retrieved from *gadm.org*.

deviants from the available household survey data, a statistical procedure was developed and tested with each of the three datasets. This way, a set of positive deviant households was identified for each country.

Qualitative follow-up interviews with positive deviants (Ethiopia, Kenya, Tanzania)

The next research step aimed at identifying 'positive deviant practices', i.e. uncommon household behaviors that are found with positive deviants and which plausibly contributed to their superior performance. These practices can be promising options for other households in similar context. In each country, between ten and 15 positive deviants were re-visited for in-depth interviews and farm observations. Three country-specific sets of positive deviant practices were identified. Guidelines for the interviews and observations were elaborated by the author of this dissertation, who carried out fieldwork in Tanzania in collaboration with Majuto Gaspar Mgimiloko (TARI-Naliendele). Interviews in Ethiopia were carried out by Selamawit Sileshi Kebede (Humboldt-Universität zu Berlin) and Jemal Nurhisen Mohammed (Mekelle University). In Kenya, fieldwork was carried out by

Jeske van de Gevel and Jerusha Onyango Achieng (Bioversity International) in collaboration with Lutheran World Relief.

Choice experiments on information preferences (Ethiopia, Kenya, Tanzania)

This dissertation explores how digital systems could use small sets of household data to select agro-advisory messages in a household-specific way. This requires linking farmers' household characteristics with their individual information needs. To generate data on smallholder farmers' information needs, individual information preferences for different advisory topics were measured through a choice experiment. The experiment consisted in ranking nine local positive deviant practices – as potential advisory topics – according to personal interest. Between 43 (Kenya) and 98 (Tanzania) of the household heads who had been surveyed earlier participated in the experiment. The choice experiment was designed by the author of this dissertation and data collection was carried out by staff of Mekelle University, Lutheran World Relief, and TARI-Naliendele. Then, using the RHOMIS data, household characteristics were identified that were useful for predicting individual information preferences. This was done by fitting a ranking model and selecting covariates that improved fit.

5 Results

5.1 The accuracy of farmer-generated data in an agricultural citizen science methodology

Agronomy for Sustainable Development (2017) 37:32

Jonathan Steinke[1,2] · Jacob van Etten[1] · Pablo Mejía Zelan[3]

1 *Bioversity International, c/o CATIE, 7170 Turrialba, Costa Rica*

2 *Humboldt-Universität zu Berlin, Division of Horticultural Economics, Unter den Linden 6, 10099 Berlin, Germany*

3 *Programa de Reconstrucción Rural, Horconcito, Santa Barbara, Honduras*

doi.org/10.1007/s13593-017-0441-y

Abstract

Over the last decades, participatory approaches involving on-farm experimentation have become more prevalent in agricultural research. Nevertheless, these approaches remain difficult to scale because they usually require close attention from well-trained professionals. Novel large-*N* participatory trials, building on recent advances in citizen science and crowdsourcing methodologies, involve large numbers of participants and little researcher supervision. Reduced supervision may affect data quality, but the "Wisdom of Crowds" principle implies that many independent observations from a diverse group of people often lead to highly accurate results when taken together. In this study, we test whether farmer-generated data in agricultural citizen science are good enough to generate valid statements about the research topic. We experimentally assess the accuracy of farmer observations in trials of crowdsourced crop variety selection that use triadic comparison of technologies (tricot). At five experimental sites in Honduras, 35 farmers (women and men) participated in tricot experiments. They ranked three varieties of common bean (*Phaseolus vulgaris* L.) with regard to *Plant vigor*, *Plant architecture*, *Pest resistance*, and *Disease resistance*. Furthermore, with a simulation approach using the empirical data, we did an order-of-magnitude estimation of the sample size of participants needed to produce relevant results. Reliability of farmers' experimental observations was generally low (Kendall's *W* 0.174 to 0.676). But aggregated observations contained information and had sufficient validity (Kendall's tau coefficient 0.33 to 0.76) to identify the correct ranking orders of varieties by fitting Mallows-Bradley-Terry models to the data. Our sample size simulation shows that low reliability can be compensated by engaging higher numbers of observers to generate statistically meaningful results, demonstrating the usefulness of the Wisdom of Crowds principle in agricultural research. In this first study on data quality from a farmer citizen science methodology, we show that realistic numbers of less than 200 participants can produce meaningful results for agricultural research by tricot-style trials.

Keywords

Citizen science · Crowdsourcing · Wisdom of Crowds · Participatory methods · Participatory Variety Selection · Common bean

1 Introduction

Agricultural research has increasingly incorporated participatory methods over the last decades in order to become more client-oriented, addressing the variable conditions and preferences of resource-poor farmers and consumers (Lilja et al. 2013). Farmer participation in agricultural research has not yet become mainstream throughout the world, however. Scaling of participation is hindered not only by institutional constraints and prejudices about data quality but also by the resource-intensity of most participatory methodologies in terms of time, training, and cost per beneficiary (Hellin et al. 2008; Morris and Bellon 2004). The scalability of current participatory approaches is often limited because they rely on organized farmer groups that require high levels of professional support and the establishment of collective plots (cf. Witcombe et al. 1996 for participatory variety selection). Moreover, these approaches require that farmers stay engaged throughout the crop cycle, even when participation involves hard work, for example, weeding the plot. This often leads to a free-rider problem. Limited participation due to free-riding has led to incomplete observations, haphazard crop variety choices, and subsequent dis-adoption of selected varieties, when farmers discover the true field performance of the variety on their own plot of land (Misiko 2013).

Alternative decentralized approaches to farmer-participatory research have been suggested recently, emphasizing linkages between farmers within research, knowledge, and innovation networks (Desclaux et al. 2008; Spielman et al. 2009). Such systems can lead to success (e.g., Humphries et al. 2015), but farmer networks require long- term commitment and cost-intensive facilitation by outsiders, and results are difficult to scale beyond the local level (Classen et al. 2008). Because increasing both empowerment and the numbers of farmers as beneficiaries are goals in participatory research (Snapp and Heong 2003), there is still an open need for methodologies that include farmer-led innovation processes *and* that are scalable.

Citizen science and "crowdsourcing" methods can help to overcome the limited scalability and free-rider problem in existing participatory methodologies for the agricultural sci- ences. Large-N citizen science projects can involve large groups of volunteers individually contributing to scientific tasks, notably data collection, by crowdsourcing approaches (Dickinson et al. 2012; Hand 2010). In these approaches, large research tasks are first subdivided into many "micro-tasks" that are doable for an individual participant. These tasks are distributed and results are collected through digital channels. The individual results are then combined to produce a large result. Crowdsourcing has been used for many applications, from translation to image recognition. When crowdsourcing involves the production of new knowledge, it relies on the "Wisdom of Crowds" principle (Surowiecki 2005). This principle implies that large groups of participants can in aggregate produce highly accurate results when certain conditions are met: a sufficient diversity of viewpoints and independence of observations. Use of information and communication technologies to receive contributions from many participants makes citizen science research scalable (Dickinson et al. 2012). As research relies on individual rather than group contributions, it also avoids the free-rider problem.

Here, we study the large-N citizen science approach proposed by Van Etten (2011) in an application to the detection of phenotypic differences between varieties of common bean, *Phaseolus vulgaris* L. (Sect. 2.1). While the full methodology involves more steps, such as participatory research priority setting, we here focus on farmer observation as a method of data collection. Large-*N* citizen science can yield accurate results if the relatively low reliability of farmers' individual observations is balanced by a large number of observations, following the Wisdom of Crowds principle. To achieve high-quality science using this citizen science approach, the level of accuracy of data collection by farmers should be clear. Farmers have heterogeneous levels of knowledge and expertise, and possess knowledge that is more developed in some domains than in others (Bentley 1989). Kremen et al. (2011) show that citizen scientists can make accurate observations in certain categories, but are more prone to bias or inaccuracy in others. Therefore, the first goal of our study is to assess the accuracy of farmers' observations in citizen science trials, as a proof of concept. Feasibility also depends on the number of participants that are needed to achieve accurate results. Therefore, the second goal of our study is to gain insights regarding the order of magnitude of the number of participants required to produce useful findings with a large-*N* citizen science methodology.

2 Materials and methods

2.1 Citizen science methodology

We apply a citizen science approach first proposed by Van Etten (2011). The approach is based on tri-adic comparisons of technologies, which we refer to here as the *tricot* approach. Tricot can be used to assess a range of agricultural technologies. When it is applied to crop varietal tests, women and men farmers each receive experimental seed

Figure 1: A farmer-managed variety selection trial for triadic comparisons of technologies (tricot) in Honduras (left). Farmers evaluating an experimental trial for the accuracy assessment reported in this article (right).

quantities of three different varieties chosen randomly from a larger set ofvarie- ties, and grow these varieties next to their own crop, under usual crop management (Fig. 1). Farmers observe the three varieties and evaluate different aspects of their performance at different points in time, using a simple ranking format, triadic comparisons. Farmers then communicate their observations to field agents verbally, on paper, or via mobile telephone. The farmer-generated observation data are

analyzed using statistical methods for ranking data. Given an adequate number of partial rankings, a preference scale for all varieties included in the experiment can be constructed by fitting a Bradley-Terry (BT) model (Bradley and Terry 1952; Coe 2002). Also, more sophisticated models for preference data can be used (e.g., Fürnkranz and Hüllermeier 2010; Strobl et al. 2011). Early experiences with applications of tricot are described by Van Etten et al. (2016).

Triadic comparisons are a proven method in ethnobiological research (Martin 2004). This format allows farmers to register and communicate their observations with a low level of literacy, and without the need to make quantitative statements. Within distinct evaluative criteria (agronomic traits, yield, processing qualities, and market value), participating farmers are asked to define each the best and the worst variety from within their set of three. Participants report their observations by answering two straightforward questions for each aspect that is being evaluated. For example, for yield, the questions would be "Which variety had the highest yield?" and "Which variety had the lowest yield?" This is an important reason for reducing the number of varieties to be tested by a single participant to three. If larger sets of varieties were to be ranked by each participant, more complicated questions would need to be asked. Straightforward questions are needed to be able to retrieve the information through telephone interviews, including automated calls.

2.2 Accuracy

We assess the quality of citizen science data produced by farmers by focusing on their accuracy. Accuracy consists of two components, reliability and validity (ISO 1994). The reliability of a method is its ability to produce repeated, consistent results. Validity refers to the closeness of a result or the mean of a large group of results to the actual value or accepted standard. The combined information about reliability and validity allows discussing the accuracy of a method.

The research method tested in this study is smallholder farmers' ranking of three different crop varieties according to observable plant characteristics. Reliability is expressed as the degree of internal agreement among observers about the ranking of varieties. Validity of the data is measured as the degree of agreement of farmers' observations with a ranking that was established by an agronomist who evaluated the same set of varieties, which we refer to as a "scientific ranking" (Sect. 2.3). The comparison between farmers' rankings and a scientific ranking is not meant to question the overall validity offarmer *knowledge*, as this would imply problematic assumptions about these two forms of knowledge and their relative value (see Cleveland and Soleri (2007) for a discussion of the epistemological questions around comparing farmer and scientific knowledge). A key motivation for participatory research is to accommodate diverse viewpoints, and to tap into knowledge that is inaccessible, hard to interpret, or “invisible” to researchers. But this means that at the same time, it is important to establish whether different elements of farmer and scientific knowledge correspond to the same objective reality. A minimal degree of correspondence is an essential condition for a meaningful dialogue between farmers and scientists.

This study has therefore the limited goal of establishing the commensurability of farmers’ and scientists’ *observations* on the same phenomena (and not their knowledge as a whole). The point is to evaluate if farmers and scientists reach the same conclusions about varietal characteristics, as a starting point for subsequent farmer-scientist dialogue to make sense of these observations. We study varietal characteristics that are objectively observable rather than characteristics that involve a strong element of subjective assessment or preference (e.g., taste).

The tricot approach makes use of the trade-off between reliability and validity by placing emphasis on validity over reliability. As the Wisdom of Crowds principle suggests, a large sample of data may lead to a correct result even when individual data entries vary strongly (low reliability) as long as an unbiased aggregate measure can be calculated from the data (high validity). Tricot achieves external validity by placing crop varieties and other agricultural technologies directly in their target environment and by evaluating their performance in the eyes of the persons who will eventually adopt the technology or not. Independence of observations is ensured by not revealing the names of varieties or technologies and asking participants individually for their results. The Wisdom of Crowds requirement of having a diversity of viewpoints is fulfilled by inviting a diverse group of participants (women, men) to grow the varieties at many different plots, each one under slightly different crop management and environmental conditions.

2.3 Experimental design

At five sites in Honduras, small trials of three different varieties of common bean (*P. vulgaris* L.) were planted by collaborating farmers. These volunteers were smallholder farmers participating in tricot-style variety selection for common bean (see Van Etten et al. 2016). We assigned to each site a combination of three different varieties drawn from a total set of seven varieties. All varieties were phenotypically clearly distinct and uniform. Seeds were obtained from the bean breeding program at Zamorano Panamerican Agricultural University in Honduras. We randomized the assignment of combinations to sites and the order in which the varieties within each combination of three were ordered. The host farmers planted the three varieties of each combination at the same date, and each farmer managed their three varieties in the same way. They located the three varieties in each set directly next to each other in sub-plots with six rows of 8 m for each variety.

At five different points in the growing cycle, a total number of 35 smallholder farmers (18 women and 17 men) were asked to evaluate the three varieties at one of the sites (Fig. 1). In each session, groups of five to eight farmers participated. The selection of participants was determined by ongoing work of two local NGOs, and no additional criteria besides a balanced gender ratio were applied. The participants were first informed about the format of the exercise and that they would be asked to evaluate four agronomic traits: *Plant vigor*, *Plant architecture*, *Pest resistance*, and *Disease resistance*. In earlier participatory research, local farmers and breeders had established these traits as the most important pre-harvest se- lection criteria for bean varietal improvement (Steinke 2015), and they are common criteria in participatory variety selection for common bean (Asfaw et al. 2012).

Participants were then asked to take a few minutes to familiarize themselves with the three varieties planted, and focus on observable expressions of the traits. From the earlier research experiences, the farmers were acquainted with the concepts of Plant vigor – a merger of leaf area, leaf color, and physiological plant state (e.g., absence of drought stress symptoms) – and Plant architecture, for which farmers prefer non-trailing, upright-growing plants. But the enumerator also rephrased the exercise using local farmers' common wording, like "how well the foliage has developed" (for Plant vigor) and "how nicely the plant stands/grows" (for Plant architecture). For pest and disease resistance, participants were asked to acknowledge the presence or absence of attack symptoms, in order to identify different resistance capacity of varieties indirectly. The rationale behind observing the occurrence of biotic stressors as an inverse proxy for resistance requires the assumption that pest and disease pressure on the three trial varieties is equal, and the intensity of attack symptoms is thus determined largely by differences in genetic resistance. The questions asked were "which variety is (least/most) affected by (pests/diseases)?"

Except for the individual host farmers, participants had not seen the trials before. The importance of independent, individual assessments was emphasized when explaining the experiment to the participants, and participants were requested to refrain from exchanging their ideas about the varieties, in order to guarantee independent observations. The participants did largely remain silent during the evaluation.

After a few minutes of observing all four traits of the three varieties, the farmers were approached individually by the enumerator. The enumerator asked for their view on which was the best and the worst variety regarding each of the four criteria and recorded the answers. In each of the sessions, a local agricultural expert, in all cases an agronomist with much field experience working with common bean, also answered the same questions, and these assessments were taken as the respective scientific ranking for each of the different sites to measure the accuracy of farmers' observations against (Sect. 2.5).

Due to differences in planting dates and growing environments, the trials were in different development stages during the fieldwork period. This limited the observations that could be made in different sessions with farmers. In particular, pest and disease incidence cannot be evaluated before plants enter the reproductive phase (approximately 35 days after sowing), so these observations were only collected at two out of five sites.

2.4 Data preprocessing

For each plant characteristic, participating farmers indicated which they found to be the best and worst out of three varieties planted in the trial, coded A, B, and C. By inserting the implicit medium-ranked variety, every individual observation was converted into a ranking pattern, for example

C > B > A. Incomplete observations and ties were removed from the data. Given the small number of observations per session, we decided to pool data from all sites by plant characteristic. For each site, farmer observations were recorded in relation to the expert's ranking order. At every site and for every evaluative criterion, the best variety according to the expert was coded variety X, the second-best variety Y, and the worst variety Z. This way, all valid farmer observations on one evaluative criterion could be converted in a standardized way to a permutation of X > Y > Z, the scientific ranking order. This way of data pooling assumes that there are no important differences between the sites in terms of the difficulty to discriminate between varieties. This is a reasonable assumption because at all sites, the local expert was able to rank the varieties for all plant characteristics in an unambiguous way (e.g., no ties between varieties); thus, any differences in rankings are mainly due to farmers' observation and interpretation ability.

2.5 Kendall's tau coefficient

To approach validity ofobservations, we quantified deviations offarmer rankings from the respective scientific ranking with Kendall's tau coefficient (τ), a measure of similarity between two rankings (Kendall 1938). The τ between two rankings is defined as follows:

$$\tau = \frac{C - D}{n(n-1)/2}$$

where C is the number of concordantly ranked item binaries (e.g., X > Y) between the two ranking lists, D is the number of discordantly ranked binaries, and n is the total number of binaries. τ may take values from –1 (completely reverse ranking) to 1 (identical ranking). In our case, the correct ranking pattern is always defined as X > Y > Z. In this case, a stated farmer observation of X > Z > Y or Y > X > Z gives $\tau = 0.33$, and Y > Z > X or Z > X > Y gives $\tau = -0.33$. Distributions of τ can be compared to the expected distribution of τ under a random null model. Under the null model, $\tau = 1$ is expected to occur in one out of six random rankings, $\tau = 0.33$ in two out of six, $\tau = -0.33$ in two out of six, and $\tau = -1$ in one out of six. To test whether there is an influence of gender on variety preferences or data quality, we performed Wilcoxon's signed rank test on the distributions of Kendall's tau coefficients of men's and women's observations for each of the four plant characteristics.

2.6 Mallows-Bradley-Terry model

For every plant characteristic, we fit a Mallows-Bradley-Terry (MBT) model (Mallows 1957; Tversky 1972) to the observed frequencies of the variety ranking patterns. Our criterion for validity was whether the MBT model was able to correctly distinguish the three varieties from each other at the $p < 0.05$ significance level. To reduce the risk of type I error due to multiple hypothesis testing, we performed *p* value corrections by the Holm-Bonferroni method (Holm 1979), a conservative method for controlling the family-wise error rate.

2.7 Kendall's *W*

We assessed reliability by determining the concordance between participants. We used Kendall's *W* to quantify the internal reliability for multiple dependent rankings (Kendall and Babington-Smith 1939). Kendall's *W* may take values ranging from 0, representing completely random results and no notice- able concordance among observers (rankers), to 1, meaning total agreement among all observers. We converted Kendall's *W* into verbal statements on agreement (from "very weak" to "unusually strong"), following the classification proposed by Schmidt (1997).

2.8 Simulations

Sample size choices will depend on trade-offs between research costs and data quality in different contexts. To inform such decisions, we created different scenarios with different numbers of varieties (n_{var}) and participants (n_{obs}). For each scenario, we determined the discriminative ability, defined as the number of varieties that can be statistically distinguished from the best variety ($p < 0.05$), as a simple heuristic.

We represent the observable performances of the varieties by a normally distributed variable, following a variation of Henrich and Boyd's (1998) simple model of environmental learning. We assume equal inter-variety intervals between varieties, and equal standard deviations (SD = 1). We estimated inter-variety interval values from the data by fitting the Thurstone-Mosteller case V (TM) model, which assumes that underlying parameters are normally distributed with an equal standard deviation of 1 (Mosteller 1951a, b). We chose the TM model for ease of interpretation because – like Henrich and Boyd's environmental learning model – the TM model uses Gaussian distributions, whereas the (Mallows-)Bradley- Terry model uses Gumbel distributions.

From the results of the TM model, we calculated the mean interval between trial varieties using the TM parameter estimates (mean of Y-X and Z-Y). This represents the mean pairwise performance difference among three varieties drawn randomly from a total pool of seven varieties. To obtain a representative inter-variety interval, we further divided the average X-Y-Z interval by 2, the mean number of intervals separating two varieties when three varieties are drawn out of a set of seven.

From the calculated performance intervals for all observed variety traits, we only retained the highest and lowest mean interval (Plant vigor and Disease resistance) for the simulations, thus testing one "easy" and one "challenging" plant characteristic. We generated 18 sets of modeled crop varieties, each containing $n_{var} \in \{3,4,5, \ldots, 20\}$ varieties. Also, we created six sets with different numbers of observers, $n_{obs} \in \{10, 20, 50, 100, 200, 500\}$. This resulted in 18 variety sets × 6 farmer sets × 2 different variety traits = 216 different scenarios.

We ran the simulations 1000 times for each of the 216 scenarios. For each run, we created a balanced experimental design. To simulate an individual participant's observation, we drew three varieties from the overall set of n_{var} varieties following the experimental design. For these three varieties, we then drew random numbers from their respective normal distributions. Subsequently, we compared these values to create a ranking. We repeated this for all n_{obs} participants in the set. We then ran the generalized Bradley-Terry-ε model on the resulting rankings (Firth 1993). This model will not break down if one variety wins or loses from all other varieties (unlike the classic BT model) and works with more than six varieties (unlike the MBT model). It is commonly used on ranking data and leads to consistent rankings (Jeon and Kim 2013). As a simple performance measure, for each of the 1000 runs, we determined the number of varieties that could be distinguished from the best variety at the $p < 0.05$ significance level, the discriminative ability. For each of the 216 scenarios, we calculated the median discriminative ability, as well as percentiles ($p = 5$ and $p = 95$).

2.9 Computational resources

For statistical analysis, we used the R programming language and environment (R Core Team 2016). We calculated Kendall's tau coefficient (Sect. 2.5) with the *Kendall* function of package *Kendall* (McLeod 2011). To fit the win counts for MBT models (Sect. 2.6) with the *glm* function (R Core Team 2016), we constructed paired comparison matrices with the *patt.design* function of package *prefmod* (Hatzinger and Dittrich 2012) and extracted *p* values with the *stars.pval* function of package *gtools* (Warnes et al. 2014). Kendall's *W* (Sect. 2.7) was calculated using the *kendall* function of package *irr* (Gamer et al. 2012). For the simulations (Sect. 2.8), we fit TM models with the *thurstone* function of package *eba* (Wickelmaier and Schmid 2004) and BT models using the functions *countsToBinomial* and *BTm* of package *BradleyTerry2* (Turner and Firth 2012). To speed up the simulations, we ran foreach loops, using the *doParallel* package (Calaway et al. 2015), and used the *plyr* package to reformat data (Wickham 2011).

3 Results and discussion

3.1 Accuracy of farmer-generated data

Table 1 presents the share of each τ value among all observations on each plant characteristic. In the case of Plant vigor, all observers fully or almost agreed with the scientific ranking. Observations on Plant architecture and Pest resistance are slightly less clear-cut, with a mean τ to the scientific ranking of about 0.5 each. Observations on Disease resistance are, on average, most divergent, with a mean τ to the scientific ranking of 0.33.

Wilcoxon's signed rank test on the τ values of men's and women's evaluations did not reveal a gender effect on observation validity for any of the plant traits at the $p < 0.05$ significance level (Table 2). The scientific literature provides evidence for gender-biased agricultural capacity, resulting from gendered household domains, such as the cultivation of different crops by women and men, gendered focus on different steps of food production and processing in the household, or contact to extension (Peterman et al. 2010; Quisumbing et al. 2014). Such gender differences may translate into different observation accuracies for different traits. In this study, however, all participants were currently engaged in cultivating bean. We would expect a stronger gender effect on agronomic

Table 1: Kendall's tau coefficient, standard deviation (SD), and Kendall's *W* of experimental farmer variety rankings

Variable	Frequency of observations with Kendall's tau coefficient (τ)				Mean τ	SD	Observers	Kendall's *W*
	τ= 1	τ= 0.33	τ= -0.33	τ= -1				
Plant vigor	64%	36%	0%	0%	0.76	0.32	22	0.676**
Plant architecture	54%	23%	19%	4%	0.51	0.60	26	0.280**
Pest resistance	46%	38%	15%	0%	0.54	0.48	13	0.337*
Disease resistance	27%	55%	9%	9%	0.33	0.57	11	0.174
Random null model	*17%*	*33%*	*33%*	*17%*	*0*	–	–	–

Percentages do not always add to 100 because of rounding. Significance values for the calculation of Kendall's *W* are as follows: $^{*}p < 0.05$, $^{**}p < 0.001$

knowledge in situations where the task division between women and men is more pronounced.

As can be seen in Table 1, correct observations with τ =1 were consistently more frequent than a random distribution would suggest, and, in return, incorrect observations with τ = −0.33 or τ = −1 were less frequent. Only for rankings on Plant architecture were observations with τ =0.33 less frequent than a random distribution would suggest. For Disease resistance, τ = 0.33 has higher frequency than τ = 1. Under the random null model, twice as many cases with τ = 0.33 are expected than with τ = 1, as two rankings are possible for τ = 0.33 (Sect. 2.5), so this does not necessarily mean that the consensus about Disease resistance does not converge to the scientific ranking. Amore synthetic approach to determine validity is to use the MBT model.

Table 3 presents the results of MBT model estimation, including Holm-Bonferroni-adjusted *p* values. For all variables, the MBT model gives the correct ranking order; i.e., the estimate differences have the correct, negative sign, and |X-Z| > |X-Y|. For Plant vigor, the MBT model not only gives the correct order but also detects significant differences between all three varieties. For Plant architecture, all variety binaries but the best to the second-best varieties can be distinguished from each other. For Pest resistance, the expert-assessed best and worst varieties can be distinguished from each other. For Disease resistance, no variety can be distinguished from an- other at the $p < 0.05$ significance level. Nonetheless, we observe that (i) in all cases, the groups of observers converged on the same order as the agronomists (X > Y > Z) and (ii) except for Disease resistance, they were able to distinguish the best from the worst variety at the $p < 0.05$ significance level. This test was based on empirical data with a small number of observations, and in the next section, we explore the

Table 2: Mean Kendall's tau coefficient (τ) and standard deviation (SD) of men's and women's observations on four plant traits, and *p* value of Wilcoxon's signed rank test between gender-disaggregated observations

Variable	Women			Men			*p* value
	Mean τ	SD	*n*	Mean τ	SD	*n*	
Plant vigor	0.74	0.33	13	0.78	0.32	9	0.841
Plant architecture	0.67	0.49	12	0.33	0.67	14	0.189
Pest resistance	0.33	0.50	7	0.78	0.32	6	0.140
Disease resistance	0.11	0.63	6	0.60	0.33	5	0.227

consequences of these findings with increased sample sizes.

We assessed Kendall's *W* as a measure of reliability for all traits (Table 1). For rankings on Plant vigor, Kendall's *W* is 0.676, a value indicating strong agreement among the observers. Rankings on Plant architecture achieve Kendall's *W* of 0.280, and rankings on Pest resistance achieve Kendall's *W* of 0.337, revealing weak agreement among observers in both cases. Rankings on Disease resistance result in Kendall's *W* of 0.174, which may be interpreted as very weak to weak agreement. Kendall's *W* was significantly higher than zero in all cases, except for Disease resistance. However, Disease resistance was the evaluative criterion for which we had the smallest sample size (n_{obs} = 11), giving it very small statistical power.

Depending on the trait, 77–100% of the observations match or nearly match the scientific ranking (τ=1 or τ=0.33), while only a 50% match would be expected if the rankings were completely random and contained no information. For all four traits, even with low numbers of observers, the MBT model ordered the varieties in the correct order, and for three traits, the model determined that the best and the worst variety per- formed significantly different from each other. So regardless of the varying levels of reliability in the data, our results were valid in all cases of our experiment.

Reliability, however, is only high for one variable, Plant vigor. This outcome relates to the expected difficulty of participants in observing the traits. Plant vigor can be assessed easily from a distance, and differences in leaf development and color intensity can be pronounced between crop varieties. Both Plant architecture and Pest resistance require some closer inspection of individual plants and leaves, which may also be somewhat more time-consuming. Lastly, the correct observation of diseases (or their absence), especially at early stages, demands more thorough scrutiny and background knowledge, including techniques of observation. Lack of training and awareness about diseases may be suggested as a reason leading to the relatively lowest validity, i.e., the highest degree of incorrect observations on Disease resistance. Our results concur with Bentley's (1989) reasoning that the ease of visual observation is an important determinant of the accuracy of farmers' observations and is therefore a main factor explaining the depth and level of concurrence of farmers' and formal scientific knowledge in different domains.

The relatively short time available for farmers' on-site evaluations in the experimental procedure we applied may also explain observed differences in accuracy to some extent. Although participants were not being rushed in our experiments, we expect that farmers in future tricot-style on-farm variety trials would get a better insight into pest and disease resistance, as farmers will be able to observe the plants on multiple occasions throughout

Table 3: Results of Mallows-Bradley-Terry model estimation of farmers' variety rankings

Variable	Varieties	Estimate difference	Standard error	z value	*p* value (unadjusted)	*p* value (Holm-Bonferroni correction)
Plant vigor	X Y	-0.895	0.293	-3.050	0.002**	0.005**
	Y Z	-0.609	0.239	-2.543	0.011*	0.011*
	X Z	-1.504	0.371	-4.049	5.152	0.000***
Plant architecture	X Y	-0.204	0.154	-1.326	0.185	0.185
	Y Z	-0.410	0.164	-2.498	0.012*	0.025*
	X Z	-0.614	0.178	-3.449	0.001**	0.002**
Pest resistance	X Y	-0.285	0.227	-1.252	0.211	0.211
	Y Z	-0.429	0.240	-1.789	0.074	0.147
	X Z	-0.713	0.270	-2.640	0.008**	0.025*
Disease resistance	X Y	-0.150	0.226	-0.663	0.507	0.507
	Y Z	-0.301	0.234	-1.283	0.199	0.399
	X Z	-0.451	0.246	-1.832	0.067	0.201

Varieties X, Y, and Z represent the expert-assessed best, second-, and third-best varieties at each experimental site, respectively.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

a growth cycle, and follow the evolution of pests and diseases over time.

Rankings that differ from the scientific ranking do not necessarily reflect a random error due to a lack of observational or diagnostic capacity, but may indicate that some participants had a semantic understanding of the concept that is divergent from the expert's understanding. For example, the concept of good plant architecture may vary among farmers, so the observers giving reverse or near-reverse rankings ($\tau = -1$ or $\tau = -0.33$) may actually have assessed correctly according to their own criteria. It may be possible to detect the presence of disagreements statistically (cf. Mueller and Veinott 2008). The detection of substantial disagreements could be used as a data quality diagnostic tool in future applications.

This study only focused on pre-harvest plant characteristics. The tricot methodology can also be employed for assessing harvest and post-harvest variety characteristics, such as yield, cooking time, and processing or storage qualities. While the findings on observation accuracy can perhaps be generalized to other vegetative plant characteristics, the experimental process described here should be repeated in order to assess the appropriateness of the tricot method for producing findings about post-harvest variables.

3.2 Discriminative ability simulations

The mean inter-variety performance interval from the TM model estimation was highest for Plant vigor (1.03) and lowest for Disease resistance (0.37). Only these values (after dividing by 2, as explained in Sect. 2.8) were used in the simulations. Figure 2 shows the median discriminative abilities, i.e., the numbers of varieties that could be distinguished from the best variety at the $p < 0.05$ significance level, as well as the respective number of varieties that could not be distinguished. In the simulation results, we observe three patterns.

Our first observation is that discriminative ability increases with an increasing number of observers. For example, the discriminative ability for Plant vigor with $n_{var} = 12$ goes up from four varieties ($n_{obs} = 10$) to ten varieties ($n_{obs} = 500$). When more observers are engaged, pairwise combinations of two varieties are replicated more often, which in turn leads to more accurate parameter estimates and a higher discriminative ability.

Secondly, we observe that discriminative ability increases when adding more varieties to the evaluation. It does so within successive sets of n_{var} values, within which every $n_{var} + 1$ leads to an increase in the number of distinguishable varieties, while the number of non-distinguishable varieties remains stable. For example, for Plant vigor and $n_{obs} = 50$, within the range of $n_{var} \in \{7, ..., 14\}$, every additional variety in the roster leads to an increase in discriminative ability. But because with every set of n_{var} values, the number of varieties that *cannot* be distinguished from the best increases by one, the relative share of distinguishable varieties among all evaluated varieties decreases overall with increasing n_{var}. For Plant vigor and $n_{obs} = 50$, for example, the share of distinguishable varieties decreases from 100% at $n_{var} = 3$ to 84% at $n_{var} = 20$. This decrease was expected; when more varieties are included in the scenario, while keeping the observer number constant, the number of evaluations of each pairwise combination of two varieties decreases. The share of distinguishable varieties may be used as a measure of efficiency of experimental design in the tricot approach.

Thirdly, for the same scenario, i.e., the same combination of n_{var} and n_{obs}, our simulations predict higher discriminative ability for Plant vigor than for Disease resistance, with few exceptions of no difference at n_{obs} =10 and n_{obs} = 500. This was to be expected, as the inter-variety intervals (relative to the standard deviation, set to 1) for Plant vigor are higher than for Disease resistance. Hence, the discriminative ability of a given number of observers will depend on the expected reliability for the tested trait, which itself depends on the ease of visual observation. For the same scenario, the discriminative ability is usually lower for Disease resistance than for Plant vigor due to the lower reliability of observations. Engaging more participants can compensate this effect. For example, our simulations indicate that the discriminative ability reached for Plant vigor with n_{var} = 12, and 50

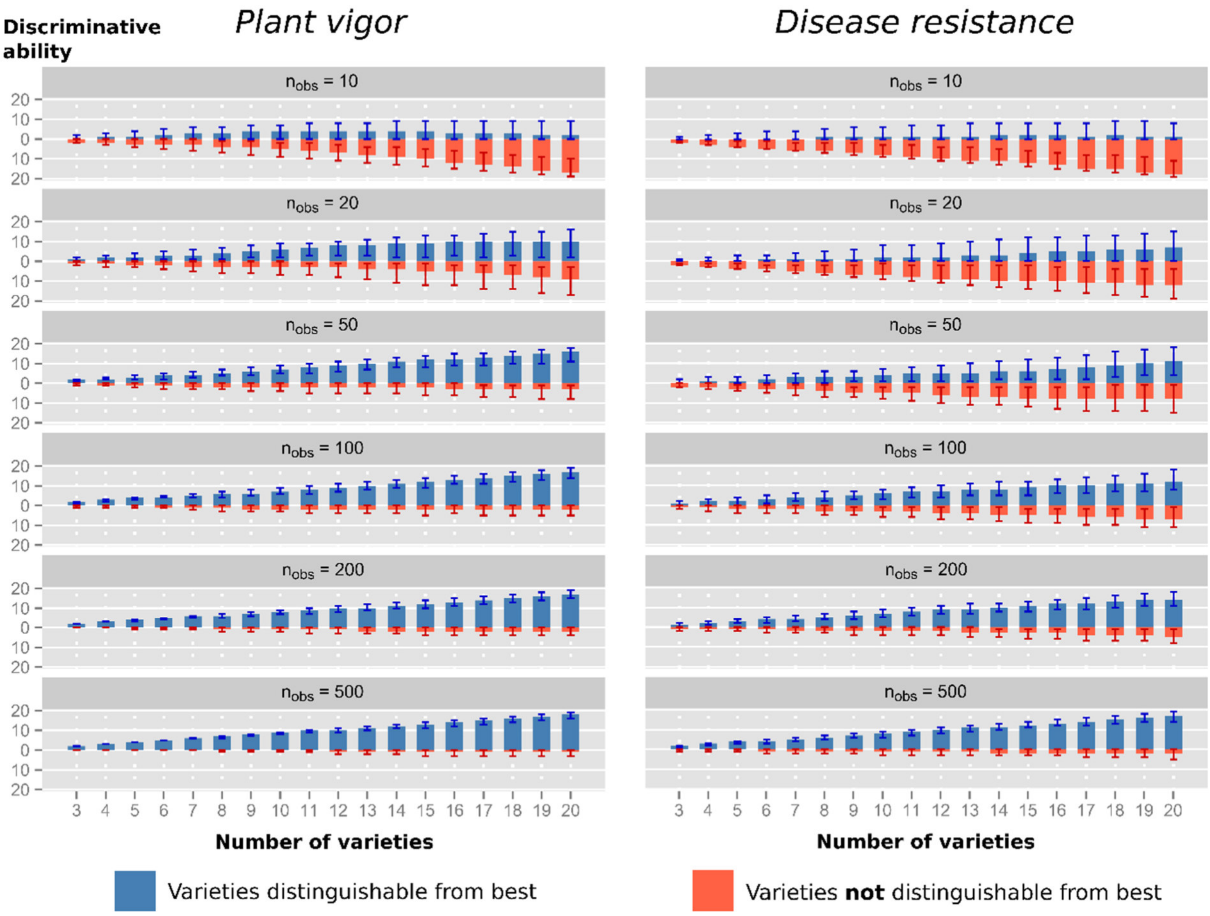

Figure 2: Simulated discriminative abilities of various research scenarios for tricot. n_{obs} = number of observers. Bars represent the number of varieties that can and cannot be distinguished from the best variety, for the "easy observation" trait Plant vigor and the "challenging observation" trait Disease resistance. Median values and percentiles (5, 95) of 1000 runs are shown.

participants would take 200 participants for Disease resistance.

How these results translate into sample size decisions will depend on the objective of variety selection. For example, a staged selection could be done, first focusing on the more easily observable characteristics. For such a first stage, only the reliability of observations on these easy traits would need to be taken into account. Also, the reliability of the observations can be increased by more training on disease recognition and other relatively challenging traits. In practical applications of the tricot approach, maximum or near-maximum discriminative ability may not be necessary. For example, the ability to identify a 50% share of varieties that perform worse than the best one may be the main aim of certain applications, e.g., to identify promising varieties at an initial on-farm screening step.

For the correct interpretation of our simulation results, it is important to note that our model assumes an idealized situation, where observable performance intervals between varieties are regularly spaced. In real life, such clear-cut differences between crop varieties are not to be expected, so discriminative ability is likely to be smaller than it is in our simulations. The selected set of varieties may include varieties that are similar for a number of traits. When the performance of varieties is virtually equal, discriminative ability may be affected. But at the same time, distinguishing between tiny differences in variety performance on farms is of limited practical relevance. Another important limitation of the current study is that it has taken into account only one potential source of error, that is, farmers' observations. Other sources of error can include experimental errors, or cases in which seeds and codes have been mixed up at some stage of the process. Also, attrition rates have not been taken into account, e.g., participants who drop out from the tricot experiment before successfully ending the trial, due to external factors or a lack of interest. To determine minimum sample sizes in real experiments, these additional factors need to be taken into account.

Furthermore, the indicated sample sizes suppose that the results are valid across the entire group of participants, which is true only in the absence of strong genotype-by-environment interactions or preferences influenced by gender, culture, or socio-economic status. Accounting for environmental gradients or doing a gender-differentiated analysis is possible, for example, by using BT models with "recursive partitioning," a method to distinguish groups of observers with different preference profiles (Strobl et al. 2011). In this case, researchers will need to revise the participant numbers upwards in order to attain reasonable results. They may use a simulation approach similar to the one presented here to assess how many participants are needed.

4 Conclusions

Our results show that in the triadic comparisons of technologies (tricot) citizen science methodology, the relatively low reliability of individual results does not undermine the accuracy of the findings when a sufficiently large group of farmers participates. Low reliability of farmer observations is no hindrance to obtaining statistically significant and relevant results. Our results show that, in aggregate, the observations contain sufficient information. Larger numbers of observations are expected to lead to statistical modeling results that distinguish between more varieties. In other words, the Wisdom of Crowds principle applies in this context: sufficiently large numbers of observers can compensate low reliability of observations as long as there is good validity, i.e., when the consensus of this large group converges on the correct answer. This means that scaling on-farm agricultural research by a crowdsourcing methodology is feasible.

Variation in farmers' observations, leading to decreased reliability, is caused not only by incorrect observations, e.g., due to the challenging evaluation of some plant traits, but also by possibly divergent views on varietal quality indicators among observers. Such differing reference systems may stem, e.g., not only from local variation in environmental pressures but also from group-specific, e.g., gendered preferences. While low reliability from either source can be balanced by engaging higher numbers of observers to achieve significant distinction of varieties, results from tricot-style research necessarily reflect an averaged approach to farmers' understandings of tested traits, as well as their possibly varying preferences. In ongoing research, we are currently testing statistical methods that treat variation as information and that lead to alternative models, disaggregating results, e.g., along cutpoints on environmental gradients.

For the varietal characteristics tested in this study, it was possible to reproduce scientific judgments through crowdsourcing farmer observations. Whether the same approach can be used to tap into farmer knowledge that is embedded in context and is inaccessible to scientists, and thereby elicit technology rankings that cannot be performed by conventional methods, remains to be tested. Our simulation results show that the order of magnitude of the group of participants required to achieve accurate results is reasonable given the logistical abilities of many organizations. Assuming an attrition rate of 20% or less, we estimate that in evaluations of sets of about 10–12 varieties, groups of 150–200 participants are likely to be sufficient to produce meaningful findings. But these results need to be revisited when more studies using the tricot approach become available. Some investment in training farmers to observe certain traits can pay off if this reduces the error significantly. Results may improve over time when farmers repeat participation over a number of crop cycles.

The possibility of citizen science via triadic comparisons of technologies opens interesting perspectives for agricultural science, beyond crop variety research. By testing technologies across environmental or socio-economic gradients, the acceptability of sets of research products can be estimated in a robust and cost-efficient way, informing the targeting of these products to certain environments and types of farms. Compared with other farmer-participatory research methodologies, adopting a "hands-off" citizen science approach reduces requirements for logistics, farmer training, field visits, and physical assets per participant. With limited resources, research organizations may reach both higher numbers and a higher diversity of farming households for the specification of technologies under development, like unreleased crop varieties. Maybe more importantly, tricot-style research can integrate new research products continuously. With every crop cycle, for example, the worst-performing fraction of the materials (varieties, lines, clones, landraces, etc.) may be exchanged with new ones. This way, through iterative research cycles, technology specification may improve, and individual participant farmers' experimentation may benefit from knowledge generated by the Wisdom of Crowds.

Recent approaches to agricultural extension have stressed the need to link stakeholders for knowledge exchange and social learning, as well as the need to facilitate autonomous experimentation with innovations (Desclaux et al. 2008; Schut et al. 2016). Steinke and Van Etten (2016) also encourage researchers employing the tricot methodology to bring together farmer citizen scientists in workshops. Yet, the benefit of the citizen science approach is that it poses a low entry threshold to those farm households who are regularly excluded from both traditional and modern approaches to extension and participatory research due to remoteness, time and labor constraints, or social conflict. Through tricot, participation in agricultural research and extension may be feasible with very low additional effort and little modification to regular farm-life activities. In addition, as observations are performed individually, under real- life farm conditions, and trait-by-trait along the crop cycle, selection will incorporate information about the variation among farmers and environments. Farmer groups working with collective plots tend to mask much of this variation (cf. Misiko 2013). Through reductions in staff time and logistics, we expect higher cost-efficiency of the approach, which we currently quantify in ongoing research. We also test the possibility of detecting the influence of environment (climate and soil) and other variables on farmer observations, and the effect of the tricot approach on farmer learning, which is an important goal of participatory research.

To researchers interested in implementing the tricot approach, we recommend to plan their research based on a preparatory order-of-magnitude study following a similar proto- col as the one presented here, as levels of discriminative ability in practice are likely to vary. A preparatory study could also detect farmers' semantic disagreement about concepts. If such disagreement is found, it can be countered by ensuring consensus about the concepts through a good explanation or by capturing the subjective element in the evaluations in a different way. Learning and exchange of experiences should iteratively help to improve the design and execution of tricot trials.

Acknowledgments

This work was implemented as part of the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), which is carried out with support from CGIAR Fund Donors and through bilateral funding agreements. For details please visit https://ccafs.cgiar.org/donors. The views expressed in this document cannot be taken to reflect the official opinions of these organizations. The project has been supported by USAID grant AID-OAA-F-14-00035 "Crowdsourcing crop improvement: Evidence base and outscaling model". The authors are very grateful to Dr. Juan Carlos Rosas and staff at Zamorano, as well as the field facilitators of the Honduran NGOs PRR and FIPAH for the support on site. We would like to thank Dr. Helen Ogden (Warwick) for an important pointer on statistics and Vincent Johnson for editing the manuscript. We thank five anonymous reviewers, whose contributions substantially improved the manuscript. Above all, we are grateful to the participating farmers for their commitment, time, and effort.

Open Access

References

Asfaw A, Almekinders CJM, Blair MW, Struik PC (2012) Participatory approach in common bean (Phaseolus vulgaris L.) breeding for drought tolerance for southern Ethiopia. *Plant Breeding* 131:125– 145

Bentley JW (1989) What farmers don't know can't help them: the strengths and weaknesses of indigenous technical knowledge in Honduras. *Agric Hum Values* Summer:25–31

Bradley RA, Terry ME (1952) Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* 39:324– 345

Calaway R, Weston S, Tenenbaum D (2015) doParallel: foreach parallel adaptor for the 'parallel' package. Available at: https://cran.r-project. org/web/packages/doParallel/index.html

Classen L, Humphries S, Fitzsimmons J, Kaaria S, Jiménez J, Sierra F, Gallardo O (2008) Opening participatory spaces for the most marginal: learning from collection action in the Honduran hillsides. *World Dev* 36:2402–2420

Cleveland D, Soleri D (2007) Farmer knowledge and scientist knowledge in sustainable agricultural development: ontology, epistemology and praxis. *In:* Sillitoe P (ed) Local science vs global science: approaches to indigenous knowledge in international development. Bergahn, New York, pp 209–230

Coe R (2002) Analyzing ranking and rating data from participatory on-farm trials. *In:* Bellon MR, Reeves J (eds) Quantitative analysis of data from participatory methods in plant breeding. CIMMYT, Mexico, D.F., pp 44–65

Desclaux D, Nolot JM, Chiffoleau Y, Gozé E, Leclerc C (2008) Changes in the concept of genotype × environment interactions to fit agriculture diversification and decentralized participatory plant breeding: pluridisciplinary point of view. *Euphytica* 163:533–546

Dickinson JL, Shirk J, Bonter D, Bonney R, Crain RL, Martin J, Phillips T, Purcell K (2012) The current state of citizen science as a tool for ecological research and public engagement. *Front Ecol Environ* 10: 291–297

Firth D (1993) Bias reduction of maximum likelihood estimates. *Biometrika* 80:27–38

Fürnkranz J, Hüllermeier E (eds) (2010) Preference learning. Springer, Berlin and Heidelberg

Gamer M, Lemon J, Fellow I. Singh P (2012) irr: various coefficients of interrater reliability and agreement. R package version 0.84, avail- able at: http://cran.r-project.org/package=irr

Hand E (2010) People power. *Nature* 466:685–687

Hatzinger R, Dittrich R (2012) prefmod: an R package for modeling preferences based on paired comparisons, rankings, or ratings. *J Stat Softw* 48:1–31

Humphries S, Rosas JC, Gómez M, Jiménez J, Sierra F, Gallardo O, Avila C, Barahona M (2015) Synergies at the interface of farmer–scientist partnerships: agricultural innovation through participatory research and plant breeding in Honduras. *Agr Food Sec* 4:1–17

Hellin J, Bellon MR, Badstue L, Dixon J, La Rovere R (2008) Increasing the impacts of participatory research. *Expl Agric* 44:81–95

Henrich J, Boyd R (1998) The evolution of conformist transmission and the emergence of between-group differences. *Evol Hum Behav* 19: 215–224

Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6:65–70

ISO (1994) Accuracy (trueness and precision) of measurement methods and results – part 1: general principles and definitions. BS ISO 5725-1. International Organization for Standardization, Geneva

Jeon JJ, Kim Y (2013) Revisiting the Bradley-Terry model and its application to information retrieval. *J Korean Dat Inf Sci* 24:1098–1099

Kendall MG (1938) A new measure of rank correlation. *Biometrika* 30: 81–93

Kendall MG, Babington-Smith B (1939) The problem of m rankings. *Ann Math Stat* 10:275–287

Kremen C, Ullman KS, Thorp RW (2011) Evaluating the quality of citizen-scientist data on pollinator communities. *Conserv Biol* 25: 607–617

Lilja N, Dixon J, Eade D (eds) (2013) Participatory research and gender analysis: new approaches. Routledge, London

Martin GJ (2004) Ethnobotany: a methods manual. Earthscan, London

Mallows CL (1957) Non-null rankingmodels. I. *Biometrika* 44:114–130

McLeod AI (2011) Kendall: Kendall rank correlation and Mann-Kendall trend test. R package version 2.2. http://CRAN.R-project.org/ package=Kendall

Misiko M (2013) Dilemma in participatory selection of varieties. *Agric Syst* 119:35–42

Morris ML, Bellon MR (2004) Participatory plant breeding research: opportunities and challenges for the international crop improvement system. *Euphytica* 136:21–35

Mosteller F (1951a) Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika* 16:3–9

Mosteller F (1951b) Remarks on the method ofpaired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. *Psychometrika* 16: 207–218

Mueller ST, Veinott ES (2008) Cultural mixture modeling: a method for identifying cultural consensus. *ARA Technol Rev* 4:39–45

Peterman A, Behrman J, Quisumbing A (2010) A review of empirical evidence on gender differences in nonland agricultural inputs, technology, and services in developing countries. *In:* Quisumbing et al. (eds) Gender in agriculture. Closing the knowledge gap. Food and Agriculture Organization of the United Nations, Washington, D.C, pp 145–186

Quisumbing AR, Meinzen-Dick R, Raney TL, Croppenstedt A, Behrman JA, Peterman A (2014) Closing the knowledge gap on gender in agriculture. *In:* Quisumbing et al (eds) Gender in agriculture. Closing the knowledge gap. Food and Agriculture Organization of the United Nations, Washington, D.C, pp 3–27

R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Schmidt RC (1997) Managing Delphi surveys using nonparametric statistical techniques. *Decis Sci* 28:763–774

Schut M, Klerckx L, Sartas M, Lamers D, Campbell MMC, Ogbonna I, Kaushik P, Atta-Krah K, Leeuwis C (2016) Innovation platforms: experiences with their institutional embedding in agricultural development. *Expl Agric* 52:537–561

Snapp S, Heong KL (2003) Scaling up and out. *In:* Pound B, Snapp S, McDougall C, Braun A (eds) Managing natural resources for sustainable livelihoods. London and Sterling, Uniting Science and Participation, International Development Research Center/Earthscan, pp 67–87

Spielman DJ, Ekboir J, Davis K (2009) The art and science of innovation systems inquiry: applications to sub-Saharan African agriculture. *Technol Soc* 31:399–405

Steinke J (2015) Citizen science with resource-poor farmers as a new approach to climate adaptation and food security: evidence from Honduras. Humboldt University Berlin, Master’s thesis

Steinke J, van Etten J (2016) Farmer experimentation for climate adaptation with triadic comparisons of technologies (tricot). A methodological guide. Bioversity International, Rome

Strobl C, Wickelmaier F, Zeileis A (2011) Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. *J Educ Behav Stat* 36:135–153

Surowiecki J (2005) The Wisdom of Crowds. Anchor Books, New York

Turner H, Firth D (2012) Bradley-Terry models in R: the BradleyTerry2 package. *J Stat Softw* 48:1–21

Tversky A (1972) Elimination by aspects: a theory of choice. *Psychol Rev* 79:281–299

Van Etten J (2011) Crowdsourcing crop improvement in sub-Saharan Africa: a proposal for a scalable and inclusive approach to food security. *IDS Bull* 42:102–110

Van Etten J, Beza E, Caldererer L, van Duijvendijk K, Fadda C, Fantahun B, Kidane YG, van den Gevel J, Gupta A, Mengistu DK, Kiambi D, Mathur PN, Mercado L, Mittra S, Mollel M, Rosas JC, Steinke J, Suchini JG, Zimmerer KS (2016) First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Expl Agric*, first View

Warnes GR, Bolker B, Lumley T (2014) gtools: various R programming tools. R package version 3.4.1. http://CRAN.R-project.org/ package=gtools

Wickelmaier F, Schmid C (2004) A Matlab function to estimate choice model parameters from paired-comparison data. *Behav Res Methods Instrum Comput* 36:29–40

Wickham H (2011) The split-apply-combine strategy for data analysis. *J Stat Softw* 40:1–29

Witcombe JR, Joshi A, Joshi KD, Sthapit BR (1996) Farmer participatory crop improvement. I. Varietal selection and breeding methods and their impact on biodiversity. *Expl Agric* 32:445–460

5.2 Prioritizing options for multi-objective agricultural development through the Positive Deviance approach

PLoS ONE (2019) 14(2):e0212926

Jonathan Steinke[1,2,3] · Majuto Gaspar Mgimiloko[4] · Frieder Graef[3] · James Hammond[5] · Mark T. van Wijk[5] · Jacob van Etten[1]

1 Bioversity International, CGIAR Research Program on Roots, Tubers and Bananas, Turrialba, Costa Rica

2 Horticultural Economics, Humboldt University Berlin, Berlin, Germany

3 Sustainable Land Use in Developing Countries, Leibniz-Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany

4 Tanzania Agricultural Research Institute, Naliendele, Tanzania

5 Livestock Systems and the Environment, International Livestock Research Institute, Nairobi, Kenya

doi.org/10.1371/journal.pone.0212926

Abstract

Agricultural development must integrate multiple objectives at the same time, including food security, income, and environmental sustainability. To help achieve these objectives, development practitioners need to prioritize concrete livelihood practices to promote to rural households. But trade-offs between objectives can lead to dilemmas in selecting practices. In addition, heterogeneity among farming households requires targeting different strategies to different types of households. Existing diversity of household resources and activities, however, may also bear solutions. We explored a new, empirical research method that identifies promising options for multi-objective development by focusing on existing cases of strong multi-dimensional household performance. The "Positive Deviance" approach signifies identifying locally viable livelihood practices from diverse households that achieve stronger performance than comparable households in the same area. These practices are promising for other local households in comparable resource contexts. The approach has been used in other domains, such as child nutrition, but has not yet been fully implemented for agricultural development with a focus on the simultaneous achievement of multiple objectives. To test our adapted version of the Positive Deviance approach, we used a quantitative survey of over 500 rural households in South-Eastern Tanzania. We identified 54 households with outstanding relative performance regarding five key development dimensions (food security, income, nutrition, environmental sustainability, and social equity). We found that, compared to other households with similar resource levels, these "positive deviants" performed strongest for food security, but only slightly better for social equity. We then re-visited a diverse sub-sample for qualitative interviews, and identified 14 uncommon, "deviant" practices that plausibly contributed to the households' superior outcomes. We illustrate how these practices can inform specific recommendations of practices for other local households in comparable resource contexts. The study demonstrates how, with the Positive Deviance approach, empirical observations of individual, outstanding households can inform discussions about locally viable agricultural development solutions in diverse household context.

Introduction

In recent years, agricultural researchers and policymakers have increasingly moved away from strategies that focus on a single goal, such as productivity or household income. Modern development paradigms, such as Sustainable Intensification [1,2] or Climate-Smart Agriculture [3] emphasize that agricultural development should pursue multiple goals at the same time, including food security, nutrition quality, and improved gender relationships. These multi-objective paradigms outline broad goals, but do not predefine interventions, though they are commonly associated with diverse practices such as agroforestry, organic farming, and farm diversification [4–6]. Choosing suitable farm-level intervention options is challenging because different contexts require different recommendations. Furthermore, trade-offs can exist between different objectives, causing dilemmas between multiple household goals [7].

To inform decision-making and design intervention strategies, various methods exist. Quantitative analysis of household data can be used for predicting the outcomes of technological and institutional change on small farms [8]. More systemic analysis considers interactions between household activities as well as trade-offs between development goals in quantitative models [9,10]. But strong complexity and systemic and behavioral uncertainties can affect the practical value of quantitative analysis for generating household-level recommendations [11]. Complementing quantitative approaches with participatory research may help to cut through this complexity and link the analysis with reality on the ground [12]. For example, to reduce the number of options to test, research has frequently subjected "best-bet" solutions to ex-ante assessments by farmers [13,14]. Participatory methods can account for context-specific considerations and preferences, but can be prone to various forms of bias, e.g., relating to the sampling of research participants [15], enumerator identity [16] or participants' resistance to modify pre-held opinions [17].

Research approaches that combine the strengths of quantitative systems analysis and participatory research to prioritize interventions are promising as they provide complementary perspectives. Existing combined approaches, however, risk underemphasizing the heterogeneity of households [11,18]. As the adoption potential of different practices can vary strongly between households, informed targeting of practices to suitable end users is required [19,20].

A combination of quantitative and qualitative methods with explicit emphasis on household heterogeneity is the *Positive Deviance approach*. This research approach was pioneered by nutritionists to identify child nutrition improvement practices that are locally viable and acceptable [21,22]. They used quantitative survey data to identify households with exceptionally good child health indicator scores compared to other households in similar circumstances. Through follow-up visits to these "positive deviants", the researchers identified feeding and hygiene practices unique to these households that possibly explained their superior performance. The identified practices were then promoted to other, worse-performing households in similar cultural and resource contexts [23]. In the field of agriculture, positive deviants have been playing key roles in innovation processes [24–26], and agricultural research has recently begun exploring systematic methods of identifying and learning from such outstanding farming households [27]. The Positive Deviance approach is an interesting data-driven approach that cuts through analytical complexity to provide suggestions on viable interventions, based on empirical, qualitative insights. Existing studies, however, did not explore smallholder household performance as a multi-dimensional phenomenon, and have not yet gone from identifying exceptionally well-performing households to identifying potentially superior practices. Our goal was to explore how the Positive Deviance approach can be adapted to identify and prioritize rural development interventions for diverse farming households that pursue multiple objectives. We describe the adapted approach, consisting of three research steps, and a case study implementation in Tanzania. Based on this experience, we discuss the potential of the Positive Deviance approach for household-specific prioritization of multi-objective development opportunities.

Methods

Overview of the approach

Step 1: In the first, quantitative research step, we collected household-level data that characterize farming systems and allow quantifying livelihood performance indicators. We used these data to identify positive deviant households that optimize household performance across multiple development objectives.

Step 2: In this qualitative research step, we explored positive deviants' behaviors through interviews and farm visits, to identify uncommon practices embedded in local context. Since alternative farming styles, involving different responses to the same trade-offs, can lead households to achieve diverse, but equally optimized farm designs [28], we expected positive deviants to employ a diverse range of practices.

Step 3: Lastly, we focused on positive deviants as success cases that can be models for other households with similar resource levels. We linked the observed practices back to the quantitative data on

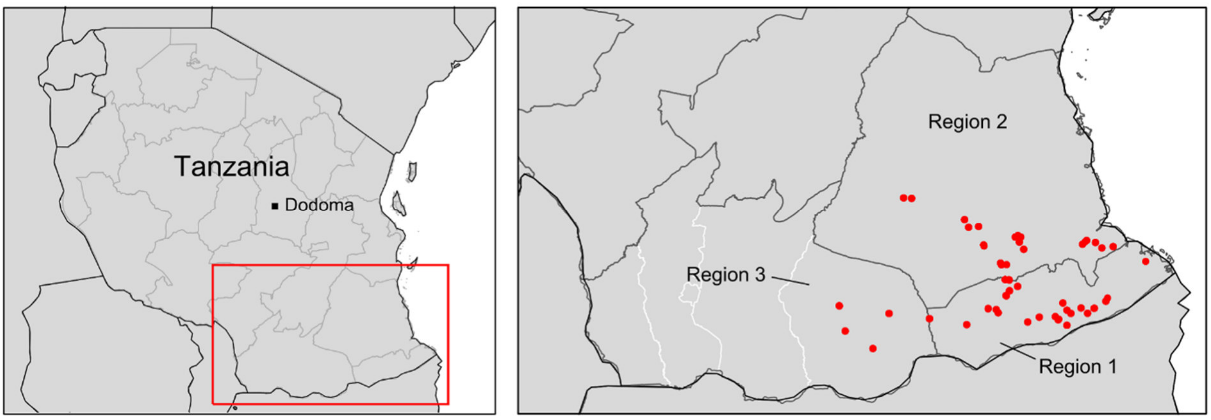

Fig 1. Research area. Household sampling sites are marked in red. Sub-regional district borders shown only where needed. Spatial data retrieved from gadm.org.

household context to estimate which practices are likely viable solutions for which particular households. We explore the feasibility of our novel method for assisting decision-making in strategic planning of development interventions, as well as providing inputs to heuristic prioritization of viable intervention options at the household level.

Research area

We conducted research in the *Southern Agricultural Zone* of Tanzania, which includes Mtwara region (Region 1, Fig 1), Lindi region (Region 2), and the Tunduru district of Ruvuma region (Region 3). Farming systems are dominated by rain-fed low-input cropping of cereals (maize, sorghum), cassava, and pulses (pigeon pea, green grams) as well as chicken husbandry for subsistence, and commercial production of pulses and oil seeds (e.g., cashew nut, groundnut, sesame). Rural population density is low (~1-5 persons/km^2), infrastructural development has been lagging behind the national standard in recent years, and poverty rates are among the highest at national scale [29].

Identification of positive deviants

Lean data household survey

We collected household data using the standardized *Rural Household Multiple Indicator Survey* (RHoMIS) [30] and calculated a set of livelihood indicators for each household (Table 1). RHoMIS provides quantitative information about individual households, including key performance variables, such as food security status and income level. It also collects data about household resources (e.g., land holdings) and the agricultural system (e.g., market orientation). To ensure data reliability, the survey collates established metrics and indicators, following standardized, replicable questionnaire formats [31–33], and reduces respondent fatigue by minimizing time burden. RHoMIS represents a snapshot view of individual households and does not aggregate or integrate information in a causal model based on “average” or “typical” household behavior.

Forty-four villages were randomly selected from administrative village lists for data collection (20 villages each in Region 1 and 2, and 4 villages in Region 3). At each village, 12 farming households were randomly sampled from lists provided by local extension officers. Two teams of four enumerators conducted the survey within a period of two weeks through face-to-face interviews at meeting points in the villages. Data was recorded and digitized on spot using the Open Data Kit software [34] on Android smartphones or tablet computers. The survey resulted in a total of 521 successful interviews with household heads.

Household performance indicators

Existing applications of the Positive Deviance approach have typically focused on single goals, such as health or nutrition. Our analysis intended to explore successful household behavior in light of possible trade-offs between different goals of current agricultural development paradigms. Despite ongoing debate, widely agreed broad goals include food security, nutrition, income, environmental sustainability, and social equity [35–37]. For each of these goals, we selected one indicator (see Table 2) and calculated household scores from RHoMIS data (see Table 1). Our choice of indicators was limited by data availability and intended to maximize

Table 1. Lean data indicators collected through the RHoMIS household survey

Indicator	Description	Unit
Household size	Household members summed up by male adult equivalent (MAE) values, accounting for different caloric energy needs and labor productivity of different gender and age groups	MAE
Household type	Marital status and gender of current household leadership. Options include: Couple, Single woman, Single man, Married woman with permanently absent spouse, Married man with permanently absent spouse	-
Land holdings	Total arable/grazing land owned by the household	Ha
Livestock holdings	Total amount of livestock, including all species, owned by the household	Tropical livestock units (TLU)
Crop diversity	Total number of different crop species cultivated during the past year	-
Livestock diversity	Total number of different livestock species owned at the moment of survey	-
Market orientation	Share of total agricultural production (in kcal) that has been sold during the past year	%
Food Availability	Potential amount of food energy generated by all on- and off-farm activities of the household, including the potential food energy bought from cash income	kcal/ MAE/ day
Number of food insecure months	Number of months the household experienced insufficient access to food of decent quality during the past year	-
Household Dietary Diversity Score (HDDS), Good Season	Number of items out of 12 different food groups (e.g., legumes, vegetables, eggs, etc.) consumed regularly by the household during the recent good season	-
Household Dietary Diversity Score (HDDS), Lean Season	See above, but during the recent lean season	-
Farm income	Total income generated through sale of farm products during the last year	US$/year
Off-farm income	Total income generated through off-farm activities during the last year	US$/year
Greenhouse gas emissions	Total on-farm greenhouse gas emissions	kg CO_2 equivalents/ year
Women's decision-making agency	Women's and female youth's cumulative share in household decision-making about benefits from on- and off-farm activities	%
Men's decision-making agency	Men's and male youth's cumulative share in household decision-making about benefits from on- and off-farm activities	%

ease of interpretation of the indicators to facilitate our analysis. Future applications may need to include more rigorous stakeholder consultation to select an agreed set of indicators.

Caloric food security. We approached food security by households' consistent access to sufficient per capita food energy, giving both consistency and sufficiency equal importance. For sufficiency, we estimated household food energy needs by multiplying household size (in male adult equivalents, MAE) by 2,550 Kcal, Tanzania's official recommended daily calorie intake per MAE [38]. The MAE concept accounts for different energy needs of household members of different genders and ages [33]. We then divided household food availability [39] by the obtained value and capped results at 100 %. For consistency, we used the number of food-secure months. We then conducted a principal component analysis on the two measures and used the first loading (which explained 57 % of variance) as a composite indicator of household food security.

Dietary diversity. Regular consumption of diverse food is crucial to a healthy nutrition. To determine household dietary diversity, we took the harmonic mean of households' HDDS scores [31] in the good and lean season, respectively (see Table 1). Unlike the arithmetic mean, harmonic mean overemphasizes lower values in the sample, generally leading to lower means. This accounted for our view that the implications to health and well-being through low nutritional diversity in one season cannot be fully balanced by a high diversity score in the other season.

Cash income. We defined disposable household cash by the sum of income from farm-gate sales and off-farm activities.

Greenhouse gas (GHG) emissions. Environmental sustainability concerns many aspects of farm management (water, soil, biodiversity) that are difficult to cover in a single indicator that would still be easy to interpret. Low farm GHG emissions are not only relevant to global climate change, which is a concern of climate-smart agriculture [3], but are also linked to agricultural practices with local environmental benefits, such as sound soil fertility

Table 2. Development goals and household performance indicators used for approximation. Indicator definitions in text

Goal	Household performance indicator
Food security	Caloric food security
Nutrition	Dietary diversity
Income	Cash income
Environmental sustainability	Greenhouse gas emissions
Social equity	Gender equity

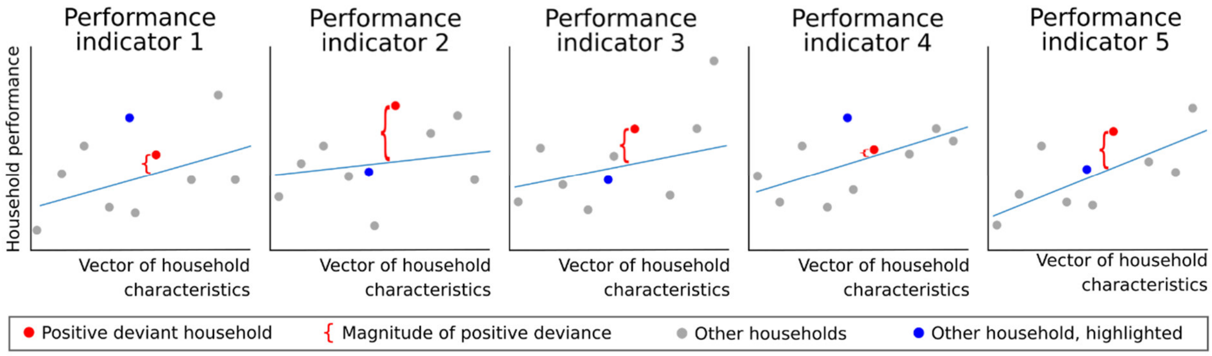

Fig 2. Conceptual figure demonstrating how performance indicators were determined from households' residuals over performance models. Light blue lines show median regressions, where performance increases with enabling household characteristics (e.g., land endowment). Positive deviants (red) are not the most successful households in absolute terms, but consistently perform better than predicted, unlike other households (see the blue dot).

management, crop rotation, and low use of chemical inputs [40]. To calculate household GHG emissions from practices reported by the households, RHoMIS uses the IPCC Tier 1 approach [32], adding up CO2-equivalents from the following emission sources and using standard emission values from literature: livestock enteric fermentation, mineral fertilizer

application, manure management, plant residue management, land use area and type, and plant-borne trace gas emissions. Because in our analysis, lower emission values imply higher sustainability, we multiplied resulting emission values by -1, resulting in increasing scores with decreasing emissions.

Gender equity. Social equity implies a fair distribution of power and benefits among many social groups, and an important societal contrast in decision-making power and benefit sharing in small-scale agriculture remains between women and men [41,42]. We therefore approach social equity by a gender equity indicator, which covers one important aspect of intra-household social equity. We calculated this proxy from the relative shares of household decision-making undertaken by women and men, respectively (see Table 1). We defined a gender-equitable situation, where decision-making is shared equally between genders, as 0.5. We then discounted deviations from the gender-equitable situation differently by household type (e.g., whether households were woman- or man-headed). The formulae are shown in the Supporting Information (S1 Table).

For each performance indicator, we capped outliers by replacing unrealistic performance scores with the maximum value observed within a realistic range. Outliers were identified by graphical plotting.

Defining and calculating deviance

We were interested in exceptional livelihood performance driven by individual household decisions and behavior. Positive deviance does not mean "a household achieves strong performance", but rather “a household’s performance is stronger than expected”. Therefore, to identify positive deviants, we transformed absolute performance into relative performance. For each dimension separately, we fit a median regression to data, using multiple household characteristics as explanatory variables to account for external determinants of performance (see below). Each household’s relative performance was thus described by the five resulting regression residuals, quantifying the difference between observed performance and performance expected based on the external determinants. We used these residuals as indicators of relative household performance (Fig 2).

As regression covariates, we used the following household variables: land endowment, livestock endowment, household size, region, and market access, all of which are known to influence livelihood outcomes [37]. Although these variables are not entirely external drivers, as they may also reflect the household’s ability in accumulating assets (land and livestock), they can be seen as constants within the scope of the intervention decisions this method is targeting. To estimate market access, we calculated the mean market orientation (see Table 1) of all households from a same village and used this average observed market utilization as a proxy for potential market access. With intra-household differences within villages evened out, we assumed that market utilization generally reflects potential market access. We eventually selected best fit performance models and included explanatory variables by the Akaike Information Criterion [43].

Pareto-optimal household performance

We defined positive deviants as households with Pareto-optimal household performance regarding the five performance indicators. Pareto-optimality does not require that positive deviants perform better than other households in each individual dimension (Figs 2 and 3). Pareto-optimal household performance means positive deviants outperform other households with equivalent characteristics in at least one dimension without being outperformed in any other dimension. This implies they optimize overall outcomes by dealing better with existing trade-offs between performance indicators. We identified positive deviants by searching for Pareto-optimal household performance in a five-dimensional space of performance scores, using the *emoa* package [44] in the R environment [45]. To obtain a reasonable number of positive deviant households in the case of our data, we ran the search twice. After the first search, we excluded the “rank 1” positive deviants from the sample and repeated the search for non-dominated households. We identified a set of “rank 2” positive deviants, which are dominated exclusively by households from the rank 1 Pareto front. In the remainder of this study, “positive deviants” refers to both groups pooled. Given the difficulty of imagining a Pareto front in a five-dimensional space, we here illustrate the concept using three dimensions (Fig 3). To create this figure, we fit a Pareto front to just three performance indicators (dietary diversity, caloric food security, cash income) in our data, and show the position of positive deviants in a three-dimensional space.

The focus on Pareto-optimality embraces diversity and does not privilege any farming style: Households that emphasize caloric food security (e.g. by intensified grain production) can be positive deviants as much as households that emphasize income generation (e.g. by value-adding). But for

Pareto-optimality, the individual performance gains must imply smaller losses in the other dimensions compared to other households, which are thus more strongly affected by trade-offs. Positive deviants with diverse priorities and activities will simply lie at different positions of the five-dimensional Pareto-front.

Households engaged in emission-intensive activities, such as cattle fattening or mineral fertilizer use, can also be positive deviants, although we use low GHG emissions as one performance indicator. Firstly, performance models consider livestock holdings, so any household's performance is always its deviation from the expected emissions level with given livestock holdings. Secondly, a positive deviant may even present high relative GHG emissions, if these do translate into increased performance in the other dimensions (e.g., generating income by value-adding dairy products, or higher crop yields).

Quantitative analysis of positive deviance

To inform strategic decision-making on interventions, we determined for which indicator and for which types of household positive deviance was strongest. We compared positive deviance both between the different dimensions of performance and along gradients of resource endowments.

To this end, we first standardized the five distributions of household performance indicators by z-transformation. Within each dimension, we subtracted the distribution mean from each score, then divided through the standard deviation. This quantified all performance scores by their distance from the mean in standard deviations, making the five indicator distributions comparable despite originally different units and scales. We then

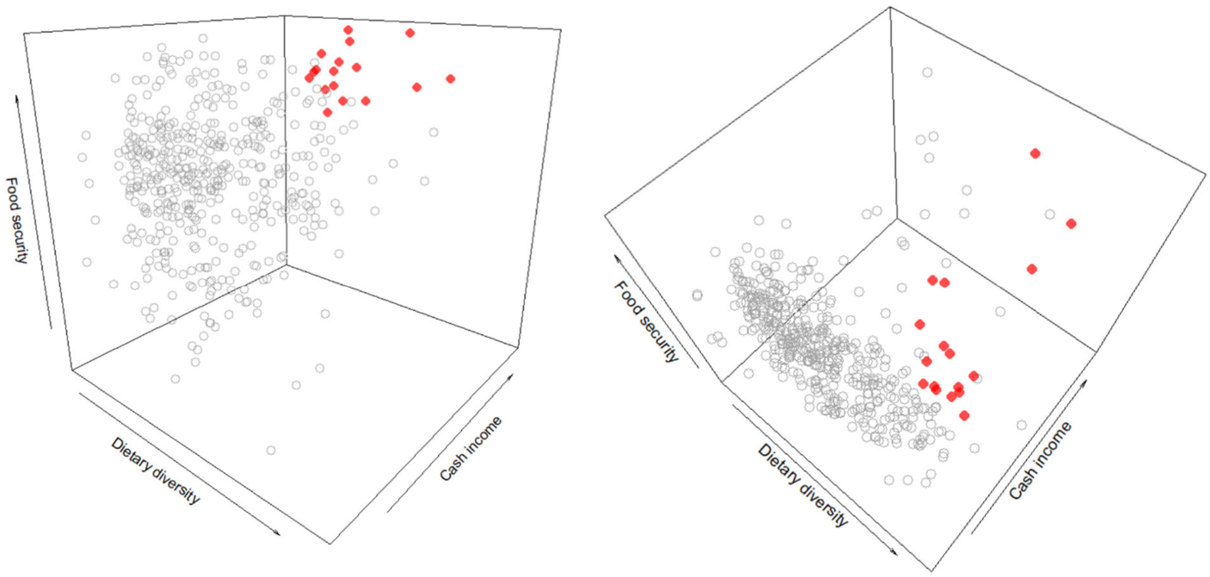

Fig 3. Location of positive deviants and other households in a three-dimensional space of household performance. Positive deviants in red, other households in grey, two perspectives on the same space. In all dimensions individually, some positive deviants are outperformed by other households, but those households suffer stronger performance losses in the respective other two dimensions.

calculated mean positive deviance of discrete subgroups of positive deviants. We defined such subgroups by household resource endowments in land and livestock. By disaggregating effects by these two key productive assets only, we intended to provide intervention agents and development planners with a simple heuristic of positive deviance in the five performance dimensions across diverse resource contexts. For this, we stratified the household sample by deciles of productive land endowments and by the median of livestock endowments (which was close to 0). The resulting 20 resource strata were thus characterized internally by similar land size and, roughly, presence or absence of livestock. We then calculated mean positive deviance of positive deviant households by performance indicator and for each resource stratum. The stratification was also used for the selection of cases for qualitative follow-up research (see next Section). To identify trade-offs between the five dimensions in realizing positive deviant outcomes, we also calculated Pearson's correlation coefficients between the magnitudes of positive deviance in the individual dimensions.

Identification of positive deviant practices

Selection of households for follow-up inquiry

Our goal was to carry out in-depth qualitative research with a diverse sub-sample of positive deviants. We selected one positive deviant household per resource stratum, applying a stepwise procedure that maximized overall diversity in household characteristics. Two of the 20 resource strata did not include any positive deviant. For the other 18 strata, we always gave preference to rank 1 positive deviants over rank 2, where rank 1 positive deviants existed. We selected the specific subset of 18 positive deviants that had highest overall diversity in terms of household size, land endowments, livestock endowments, and market access. This was the set of 18 households with maximum mean crowding distance [46] regarding those four characteristics (we excluded region, a categorical variable).

Interviews and farm visits

Of the 18 households we selected as case studies for more in-depth exploration of livelihood choices, we were able to meet 15 household heads in 12 villages. They were the same persons who had responded the lean data household survey. With every respondent, we first carried out an exploratory, semi-structured interview about the household's activities (1-3 hours), and then visited at least one farming plot together. We intended to capture all activities related to food production, storage, processing, consumption, income generation, natural resource management, and access to information, paying special attention to any details that seemed unusual (interview guideline in Supporting Information, S1 Text).

The objective of the interviews and farm visits was to identify any practices that were uncommon among most rural households and thus plausible explanations for the positive deviants' superior performance. During the interviews, we asked follow-up questions about any activities that seemed outstanding at first view. To decide which household practices were indeed uncommon, we relied on three strategies: Firstly, we also interviewed three household heads in the research region who had not participated in the lean data survey. Though we cannot determine whether they would have been positive deviants or not, we treated them as non-positive deviants. Secondly, we relied on our own experience in local farming context (especially author MGM, who participated in all interviews). Thirdly, we asked the positive deviant farmers, who often cited travels, recommendations from friends or extension agents, or personal creativity as inspiration for engaging in uncommon practices. Irrespective of the source of knowledge, we regarded as positive deviant practices all livelihood-related practices that were both uncommon in the research region and established beyond experimental stage at the positive deviant household. In joint deliberations, the authors who carried out the interviews (JS and MGM) analyzed interview notes to decide which household activities fulfilled these criteria, leading to an agreed list of observed positive deviant practices.

Positive deviants as models for similar households

In prioritizing development options for target households, we intended to account for household diversity by suggesting multiple intervention options according to individual household characteristics. We tried to avoid both over-targeting of practices (closed to households' diverse preferences) and under-targeting (letting all households choose from the full set of options). To provide a useful heuristic tool to development agents, we here focused, for each target household, on the practices found with the three positive deviants that were most similar to it. We suggest this limited number of positive deviants, along with the set of practices found with them, should inform focused discussions about viable, individually suitable development narratives grounded in local reality, through "case-based reasoning" [47].

We approached similarity between target households and positive deviants by their household endowments in six key resources: agro-ecological ability, labor, financial capital, land holdings, livestock holdings, and social capital (proxy definitions, based on RHoMIS indicators, in Supporting Information, S2 Table). For each household included in the baseline survey, we identified the three most similar positive deviants from the subsample we had visited (see previous section) by calculating Euclidean distance on the six resource levels. We defined for each of these target households the three positive deviants with lowest Euclidean distances (its 1st, 2nd, and 3rd "resource homologues"). Euclidean distance treats positive and negative deviations (whether the household's resource levels were higher or lower than those of the positive deviant) equally, accounting for some fluidity and compensation effects between resources (e.g., livestock and capital are often mutually convertible to certain extent).

Ethics statement

This study conforms with the principles of the 1964 WMA declaration of Helsinki. Approval for survey data collection was obtained from both project leadership at Bioversity International and the directorate of Naliendele Agricultural Research Institute. Research permissions for the RHoMIS survey and positive deviant interviews were also obtained from District Agricultural, Irrigation and Cooperative Officers (DAICOs) in all administrative districts included, conforming with the requirements of the Tanzania Commission for Science and Technology (COSTECH). The ethics committee at the Faculty of Life Sciences at Humboldt University Berlin was not involved because its guidelines do

Table 3. Selected socio-economic characteristics and median performance scores of surveyed households

	Positive deviants	**Other households**
Number of households	54	476
In region 1 / 2 / 3	**59 % / 26 % / 15 %**	**43 % / 48 % / 9 %**
Woman-headed households	30 %	29 %
Mean age of household leader	44.4	47.9
Education of household leader:		
Illiterate / Literate / Primary / Secondary	**2 % / 4 % / 76 % / 19 %**	**8 % / 7 % / 80 % / 5 %**
Marital status: Married	91 %	86 %
Mean household size (MAE)	4.34	4.21
Mean land endowment (Ha)	4.09	3.89
Mean livestock holdings (TLU)	0.28	0.36
Mean livestock diversity	**1.06**	**0.79**
Mean crop diversity	4.26	3.96
Presence of off-farm income	43 %	30 %
Median caloric food security (unitless)	**0.67**	**0.23**
Median dietary diversity (food groups)	**6.56**	**4.00**
Median cash income (US$/year)	686	281
Median GHG emissions (CO_2-eq/year)	**395**	**212**
Median gender equity (%)	0.33	0.33

Significant differences ($p < .05$) in household characteristics are shown bold (Student's t-test / Pearson's Chi square test).

not require prior ethical approval for a household survey like this. Survey participants were not particularly vulnerable, data was processed in anonymized form, and survey participants had the possibility to skip questions. Explicit oral informed consent was obtained from all survey participants prior to survey enumeration and documented as opening question in the RHoMIS survey. If consent was denied, enumeration stopped after one question. Permission for obtaining oral rather than written consent from survey respondents was granted by

DAICOs, given literacy limitations among the target population.

Results

Characteristics of positive deviants

Out of the 521 surveyed households, 54 were positive deviants, achieving rank 1 (n=12) or rank 2 (n=42) Pareto-optimal performance for five dimensions of household performance. Positive deviants stood out due to their strong relative performance considering their specific household characteristics. Nonetheless, for three dimensions (caloric food security, dietary diversity, and cash income), positive deviants on average also achieved higher absolute performance than other households. Overall, they did not realize higher gender equity than other households, and even showed slightly worse indicator values for GHG emissions in absolute terms (Table 3).

Positive deviants did not differ from other households with respect to gender ratio, age, marital status, household size, land endowment, and livestock endowment (Table 3). Positive deviants had,

Table 4. Mean deviance by performance dimension and aggregated resource strata

	Caloric food security[a]	Cash income (US$/a)	Dietary diversity (food groups)	Gender equity (%)	GHG emissions (CO_2-eq/a)[b]	n
Land size strata						
1+2	0.79	986	2.6	1	379	13
3+4	0.56	251	1.4	2	722	9
5+6	0.83	592	1.6	-9	834	7
7+8	0.60	461	1.9	1	359	13
9+10	0.70	3140	2.7	-5	479	10
Low livestock	1.01	1251	3.4	-5	-285	15
High livestock	0.56	994	1.6	-2	813	39
Overall mean	0.69	1066	2.1	-1	508	54
Overall mean (scaled, unitless) [c]	0.65	0.07	-0.93	-1.52	0.33	54

[a] Caloric food security scores are products of a principal component analysis and unitless.

[b] Values refer to reductions against expected values, so high values are desirable.

[c] To allow comparison of deviance across dimensions of performance, means were also scaled by z-transformation (last row). For each dimension, the unitless value quantifies mean deviance by the difference from the population mean in standard deviations.

Table 5. Deviance of individual positive deviants that were visited for qualitative follow-up research, practices identified with them, and numbers of resource homologue households per positive deviant

Positive deviant (interviewed)	Magnitude of deviance					Practices[b]	Number of resource homologue households[c]			
	Caloric Food security (unitless)	Cash income (US$/a)	Dietary diversity (food groups)	Gender equity (%)	GHG emissions (CO_2-eq/a)[a]		1st	2nd	3rd	*Total*
I	0.74	202	1.35	14	491	Sc	253	53	41	*347*
II	1.35	826	0.28	4	812	Ic	1	53	23	*77*
III	0.10	698	-1.71	1	4,492	Mb, Pi, Sc	5	7	56	*68*
IV	0.62	539	4.07	0	-161	Sc	8	10	3	*21*
V	0.29	127	3.38	1	1,618	Lb, Mt, Ss	31	56	2	*89*
VI	0.56	331	3.72	0	-103	Lb, Sc, Wl	5	7	6	*18*
VII	0.01	113	0.30	0	2,585	Pu, Tn, Sc	4	31	61	*96*
VIII	1.60	129	-0.13	1	662	Wl	59	267	0	*326*
IX[d]	1.70	10,477	2.00	1	2,338	-	-	-	-	-
X	1.35	2,081	2.73	1	-539	Lb, Mt, Pu, Wl, Tb	54	0	1	55
XI	1.49	1,819	3.61	4	-633	Cs	9	7	6	22
XII	0.00	649	4.34	2	676	Cs, Sp	25	21	8	54
XIII	1.12	513	2.56	-4	-349	Sc, Ss, Wl	52	7	289	*348*
XIV	1.54	964	3.57	1	-274	Cp, Mt	15	2	25	*42*

[a] Values refer to reductions against expected values, so high values are desirable.

[b] See Table 6

[c] As most (1st), second-most (2nd) and third-most homologue (3rd)

[d] No deviant practice identified

however, achieved higher levels of formal education. They were also not evenly distributed across regions, with significantly fewer positive deviants in Region 2 than in the other two regions. Both

Table 6. Pearson's correlation coefficients between dimension-specific magnitudes of positive deviance

	Caloric food security	Cash income	Dietary diversity	Gender equity	GHG emissions
GHG emissions	-0.22	0.24	-0.18	-0.16	1
Gender equity	**-0.29**	**-0.35**	-0.19	1	
Dietary diversity	0.26	0.20	1		
Cash income	**0.32**	1			
Caloric food security	1				

Significant relationships ($p < .05$) are shown bold.

positive deviants and other households had relatively low mean livestock endowments. Mean livestock diversity, however, was higher for positive deviants than for other households.

Overall patterns in positive deviance

Overall, mean positive deviance was strongest for caloric food security, followed by GHG emissions and cash income (Table 4, last row). For gender equity, positive deviants on average actually performed slightly weaker than expected (Table 4, last but one row). Individual positive deviants achieved diverse outcomes regarding the specific magnitudes of positive deviance in each dimension (see the examples in Table 5), and there were both weak positive and weak negative correlations between these magnitudes (Table 6).

Both land and livestock endowments seemed to influence average positive deviance (Table 4). For the smallest and largest farm sizes, positive deviance was strongest for cash income and dietary diversity. For GHG emissions, however, medium-sized farms showed strongest deviance. Household with low livestock endowments had, on average, stronger positive deviance for caloric food security, cash income, and dietary diversity. In turn, households with higher livestock endowments performed more strongly for gender equity and GHG emissions.

Positive deviant practices

Through interviews and farm observations with a subset of 15 positive deviants, we identified 14 "positive deviant" practices (Table 7 and Fig 4). We found seven of these practices with single positive deviants only, but other practices were applied by up to six positive deviants. At one household, we did not identify any uncommon practice. Other positive deviant households were, on average, engaged in 2.2 of the practices, simultaneously (maximum: 5).

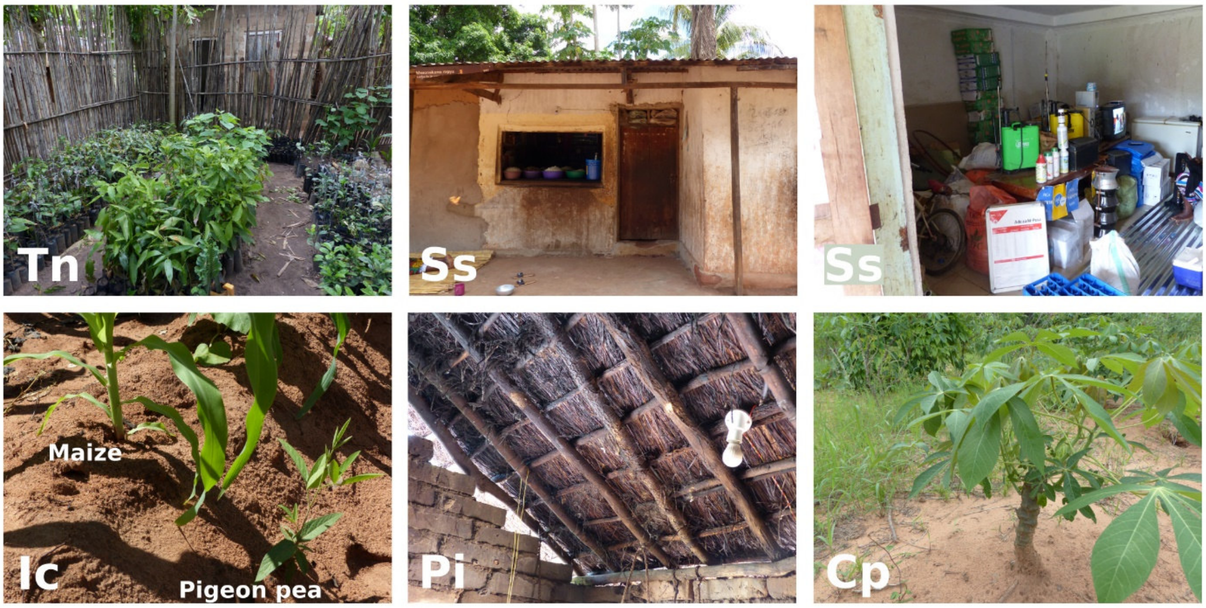

Fig 4. Examples of deviant practices observed with positive deviants. Tn, tree nursery; Ss, small shop; Ic, resource-efficient intercropping of maize and pigeon pea; Pi, poultry intensification; Cp, production of cassava planting material

Resource homologues

For each household, three positive deviants were identified according to their relative similarity to the household in resource endowments ("resource homologues"). For 323 households (62 %), the homologues were, in varying orders, positive deviants I, VIII, and XIII (see Table 5). For these households, priority interventions might emphasize farm labor scheduling (Sc) and off-farm income generation through a small shop (Ss) or wage labor (Wl). The shares of households associated to each individual practice by the resource homologue approach ranged from 8 % for the production of cassava planting material, to 100 % for farm labor scheduling (Table 7).

Discussion

Diverse positive deviants may inform household-specific intervention choices for heterogeneous target households

We designed and tested a method to identify farming households that achieve unexpectedly strong performance (positive deviants) and identified diverse practices that may have contributed to their superior outcomes. Positive deviants, about 10 % of the survey sample, represented the overall household diversity well, including, e.g., very small and very large farm sizes. Uncommon practices were found even among the least wealthy households, implying that positive deviants indeed made superior household decisions, instead of just overstating performance in the baseline survey. Regional

Table 7. Positive deviant practices observed with positive deviant households and total numbers of households that would be targeted with each practice, following the resource homologue approach (n_{max} = 521)

Practice	Code	Mechanism	Frequency observed	Number of target households	% of total
Production of cassava planting material	Cp	Generating income by producing and selling quality cutlings of an improved cassava variety	1	42	8
Investments into improved crop storage	Cs	Decreasing post-harvest losses by investing into improved crop storage constructions or triple layer PICS sacks [48]	2	76	15
Resource-efficient intercropping of maize and pigeon-pea	Ic	Decreasing plant competition for environmental resources by sowing pigeon pea at the lower end of the shadow-side slope of ridges	1	77	15
"Livestock bank"	Lb	Increasing household resilience by maintaining ruminant livestock even against short-term utility logic, for sale in emergency situations	3	107	21
Milk business	Mb	Generating income by pooling small-scale cow milk production with neighbors and sending bulk produce to buyer in town via public transport	1	68	13
Shared use of mechanical tillage	Mt	Increasing economic farm efficiency by pooling capital with neighbors to hire a tractor-tillage service provider, saving wages for manual tillage laborers	3	131	25
Intensified poultry production by artificial lighting	Pi	Increasing poultry production per unit of time by investing into a solar power-driven light bulb, enforcing artificial lighting all night and increasing daily food intake of poultry	1	68	13
Up-scaled poultry production	Pu	Increasing production and productivity of poultry by investing into bigger, more secure coops and/or new animals of improved breeds	2	96	18
Meticulous scheduling of labor allocation during land preparation and sowing of crops	Sc	Decreasing risk of crop failure by applying agronomic knowledge and skills in proper priority-setting for time and labor allocation during early phases of the growing season	6	521	100
Speculative purchase and stockpiling of crop	Sp	Generating income by investing into buying crop when prices are low, renting storage space, and selling when prices are high	1	54	10

Table 7. *continued*

Small shop for ago-inputs and building materials	Ss	Generating income by running a small village shop, often employing family members, selling agro-inputs sometimes on a commission base	1	348	67
Transportation business	Tb	Generating income by investing into a van that connects two urban centers multiple times per day, with a family member employed as driver	1	55	11
Commercial tree nursery	Tn	Generating income by producing and selling tree seedlings, including grafted cashew seedlings	1	96	18
Wage labor	Wl	Generating income by dedicating labor to off-farm wage work	4	421	81

imbalance in the distribution of positive deviants may be due to different intensities of trade-offs at different locations, e.g. due to distinct dominating farming systems. The higher livestock diversity that was observed with positive deviants might in itself represent a positive deviant practice, since livestock diversification is associated with multiple livelihood indicators [49]. That positive deviants on the whole have received higher levels of formal education is not surprising, as education is known to drive on-farm innovation processes, especially by reducing risk aversion [50], and may give farmers more lucrative off-farm labor opportunities.

The diversity in resource context among positive deviants suggests that household performance heterogeneity is at least partly due to individual decisions and behaviors. It also implies that for most households, positive deviants in relatable household context (with similar productive resources, location, farming system) may exist. This heterogeneity of success cases could be exploited to accelerate local development: For any household, the resource homologue approach identifies positive deviants as most similar solution templates, which may serve as starting points for empirically grounded discussions around adaptations in farm decision-making. This provides development agents with a heuristic for household-specific prioritization of intervention options, rather than assigning households to broad clusters, which may mask important parts of heterogeneity [51]. Since the group of positive deviants was highly diverse, such discussions may take the heterogeneity of target households into account. Given the empirical nature of insights from our method, kickstarting practitioners' discussions about interventions may require less assumptions than alternative methods that assess the effects of new practices based on household data [8]. This empirical focus, however, restricts analysis to practices that are already in use in the study area, meaning that some promising technology options, as well as institutional change, may be left out of discussions.

Identifying locally viable practices for agricultural development does not require complex econometric or system modeling

Studying the identified household success cases should allow development agents to draw plausible links between unique practices and performance outcomes. This does not require a comprehensive inventory of household activities, data-intensive system modelling or more complex econometric analysis. The method can be used by development agencies, such as NGOs or extension services, to rapidly identify a list of candidate practices that can then feed into empirically grounded discussions on intervention priorities. While the first, quantitative step requires knowledge on data cleaning and statistical analyses, it can be carried out by remote collaborators, e.g., researchers. For the second, qualitative step, the focus on empirical success cases instead of causalities, data means, and trends likely makes it easier for stakeholders not familiar with quantitative methods to participate meaningfully in discussions about viable development strategies.

Interestingly, the 14 positive deviant practices identified in this study differed from what has previously been suggested as "best-bet" solutions in similar context, such as rainwater harvesting, or biochar utilization [51]. Visiting more positive deviants and repeating the inquiry at another time of the year likely would have led to more practices, and possibly a larger overlap with the practices presented in the literature. Including a different number of households in the quantitative survey might have led to different sets of positive deviants and associated practices. The same is true for alternative indicator definitions, as we used available data from the RHoMIS survey, which provides a rapid, but also necessarily limited view of household performance. Defining performance indicators differently would likely have identified a different set of positive deviants, possibly with different practices.

More importantly, however, we identified concrete local realizations of certain practices (e.g., "resource-efficient maize-pigeon pea intercropping"), while many prioritization exercises describe broad collections of practices (e.g., "intercropping" without specifying the crops) [20,52]. The concrete practices we identified may be more directly applicable for other households. Promoting these directly observable cases may inspire others to test these practices on their own farms [53]. This can lead to further formal and informal adaptation and experimentation, perhaps supported by systematic on-farm experimentation formats [54,55].

In suggesting interventions, development agents should mind some important limitations to the effects that the identified practices can have on household performance. For example, finite societal demand for some of the produced goods and services (e.g., tree seedlings, village shops) may cap the total numbers of adopting households that may sustainably improve their livelihoods. As expected, practices that likely involve market competition (Cp, Sp, Ss, Tb, and Tn) in general seem less widely applicable than other practices, following the resource homologue approach. Potential negative societal externalities of some of the identified practices also deserve attention. For example, speculative stockpiling of crop after harvest may increase consumer prices and aggravate food insecurity of landless people. Likewise, replacing manual tillage by renting a tractor can reduce income opportunities for low-skilled, often landless rural people.

Performance differences between positive deviants and other households suggest locally promising intervention domains

Positive deviants demonstrate that household performance can be improved in each of the five dimensions. Nonetheless, there are important differences that may inform decision-making on interventions and research. For example, positive deviants on average performed considerably better than other households regarding caloric food security, but positive deviance was relatively weak for gender equity. Those households that stood out particularly for their gender equity tended to have below-average positive deviance for the other dimensions, and vice versa. While there seems to be strong potential for interventions that target food security, this trade-off indicated that less opportunities exist for improvements in gender equity without affecting other indicators negatively.

The difference may, however, also reflect current priorities of households (more experimentation around production than around social relationships) or mean that progress in gender equity requires more radical innovation, which may be less likely to develop through farmers' own experimentation [56]. Follow-up research could explore possible solutions. But future applications of the Positive Deviance approach might also reach different conclusions by using more comprehensive conceptualizations of gender equity, as we used a relatively narrow perspective on intra-household responsibilities. In addition to partial conceptualizations, our choice of household performance criteria, which was based on current development paradigms, may risk identifying success cases that are not preferred by local stakeholders. More participatory agenda-setting could be used to increase impacts in future uses of our method.

Positive Deviance constitutes a distinct, complementary approach to other participatory approaches in agricultural research. Other qualitative research approaches are also able to generate concrete example cases [57], but our method is unique in applying a highly systematic procedure with objective criteria to select a diverse subset of well-performing households. A step-wise research procedure of inquiry enhances the reliability and replicability of our method: Although the use of farmer self-reported quantitative data can introduce new forms of bias [58], the subsequent qualitative research step filters out low-quality data, as the farm visits allowed us to distinguish actual positive deviants from households that might have over-reported performance. Also, sampling diverse example cases from a reasonably large group of positive deviants (~10 % of all households) helped to avoid a narrow focus on the most extreme outliers, which may suffer more from low data quality (due to exaggeration or data entry mistakes). This principled approach likely reduced certain types of bias reported in participatory research due to less systematic selection of households and data processing [16,59]. Compared to other participatory approaches, however, our method requires an investment into prior survey data collection. Even so, in projects that require quantitative impact assessment, the RHoMIS survey can serve both as baseline and as input to the analysis of Positive Deviance.

Conclusions

We designed a new method for informed planning of household-level smallholder agricultural development interventions by operationalizing the Positive Deviance approach. A novelty in our application of the approach is the simultaneous focus on multiple objectives in agricultural development, based on the concept of Pareto-optimality. We explored how cases of surprisingly strong multi-objective household performance (positive deviants) can be identified from survey data, and how the diversity in the dataset can be exploited to inform the household-specific prioritization of intervention options for heterogeneous target households. Our analysis explored the differences between positive deviants and other households, generating a list of household-level development options that were proven to work in local context. This type of empirical insights provides valuable inputs to discussions by development practitioners and farmers for planning development interventions that are well-grounded in local context as well as conscious of trade-offs between multiple objectives. In the future, our method may be extended to other use contexts (beyond agriculture) that imply trade-offs between different development goals.

Acknowledgments

We thank Dr. Omari Mponda and the field staff at TARI-Naliendele, who collected the survey data. Our sincere gratitude goes to all participating farmers for sharing their information and showing us their farms. We thank the editor and two reviewers for their valuable comments, as well as handling staff at PLOS ONE for smooth article processing.

References

1. Garnett T, Appleby MC, Balmford A, Bateman IJ, Benton TG, Bloomer P, et al. Sustainable intensification in agriculture: Premises and policies. Science. 2013;341(6141):33–34.
2. Rockström J, Williams J, Daily G, Noble A, Matthews N, Gordon L, et al. Sustainable intensification of agriculture for human prosperity and global sustainability. Ambio. 2017;46:4–17.
3. Lipper L, Thornton P, Campbell BM, Baedeker T, Braimoh A, Bwalya M, et al. Climate-smart agriculture for food security. Nat Clim Chang. 2014;4(12):1068–1072.
4. Pretty J. Agricultural sustainability: Concepts, principles and evidence. Philos Trans R Soc B. 2008;363(1491):447–465.
5. Godfray HCJ, Garnett T. Food security and sustainable intensification. Philos Trans R Soc B. 2014;369(20120273).
6. Pretty J, Toulmin C, Williams S. Sustainable intensification in African agriculture. Int J Agric Sustain. 2011;9(1):5–24.
7. Klapwijk CJ, van Wijk MT, Rosenstock TS, van Asten PJA, Thornton PK, Giller KE. Trade-offs in agricultural systems: Current status and way forward. Curr Opin Environ Sustain. 2014;6:110–115.
8. Noltze M, Schwarze S, Qaim M. Understanding the adoption of system technologies in smallholder agriculture: The system of rice intensification (SRI) in Timor Leste. Agric Syst. 2012;108:64–73.
9. Hammond RA, Dubé L. A systems science perspective and transdisciplinary models for food and nutrition security. PNAS. 2012;109(31):12356–12363.
10. Schreinemachers P, Berger T. An agent-based simulation model of human-environment interactions in agricultural systems. Environ Model Softw. 2011;26(7):845–859.
11. Woodward SJR, Romera AJ, Beskow WB, Lovatt SJ. Better simulation modelling to support farming systems innovation: Review and synthesis. New Zeal J Agric Res. 2008;51(3):235–252.
12. Carberry P, Gladwin C, Twomlow S. Linking simulation modelling to participatory research in smallholder farming systems. In: Delve RJ, Probert ME, editors. Modelling nutrient management in tropical cropping systems. ACIAR Proceedings 114; 2004. pp. 32–46.
13. Khatri-Chhetri A, Aggarwal PK, Joshi PK, Vyas S. Farmers' prioritization of climate-smart agriculture (CSA) technologies. Agric Syst. 2017;151:184–191.
14. Schindler J, Graef F, König HJ. Participatory impact assessment: Bridging the gap between scientists' theory and farmers' practice. Agric Syst. 2016;148:38–43.
15. Cornwall A. Whose voices? Whose choices? Reflections on gender and participatory development. World Dev. 2003;31(8):1325–1342.
16. Humphreys M, Masters WA, Sandbu ME. The role of leaders in democratic deliberations: Results from a field experiment in São Tome and Principe. World Polit. 2006;58(4):583–622.
17. Stasser G, Titus W. Pooling of unshared information in group decision making: Biased information sampling during discussion. J Pers Soc Psychol. 1985;48(6):1467–1478.
18. Van Ginkel M, Sayer J, Sinclair F, Aw-Hassan A, Bossio D, Craufurd P, et al. An integrated agroecosystem and livelihood systems approach for the poor and vulnerable in dry areas. Food Secur. 2013;5:751–767.

19. Corbeels M, De Graaff J, Ndah TH, Penot E, Baudron F, Naudin K, et al. Understanding the impact and adoption of conservation agriculture in Africa: A multi-scale analysis. Agric Ecosyst Environ. 2013;187(1):155–170.

20. Mwongera C, Shikuku KM, Twyman J, Läderach P, Ampaire E, Van Asten P, et al. Climate smart agriculture rapid appraisal (CSA-RA): A tool for prioritizing context-specific climate smart agriculture technologies. Agric Syst. 2017;151:192–203.

21. Bisits Bullen PA. The positive deviance/hearth approach to reducing child malnutrition: Systematic review. Trop Med Int Heal. 2011;16(11):1354–1366.

22. Sternin M, Sternin J, Marsh D. Designing a community-based nutrition program using the Hearth model and the Positive Deviance approach – a field guide. Save the Children. Westport, CT; 1998.

23. Marsh DR, Pachón H, Schroeder DG, Ha TT, Dearden K, Lang TT, et al. Design of a prospective, randomized evaluation of an integrated nutrition program in rural Viet Nam. Food Nutr Bull. 2002;23(4):34–44.

24. Ochieng CMO. Development through Positive Deviance and its implications for economic policy making and public administration in Africa: The case of Kenyan agricultural development, 1930-2005. World Dev. 2007;35(3):454–479.

25. Pant LP, Hambly Odame H. The promise of positive deviants: bridging divides between scientific research and local practices in smallholder agriculture. Knowl Manag Dev J. 2009;5(2):160–172.

26. Biggs S. Learning from the positive to reduce rural poverty and increase social justice: Institutional innovations in agricultural and natural resources research and development. Exp Agric. 2008;44(1):37–60.

27. Modernel P, Dogliotti S, Alvarez S, Corbeels M, Picasso V, Tittonell P, et al. Identification of beef production farms in the Pampas and Campos area that stand out in economic and environmental performance. Ecol Indic. 2018;89:755–770.

28. Groot JCJ, Oomen GJM, Rossing WAH. Multi-objective optimization and design of farming systems. Agric Syst. 2012;110:63–77.

29. Perfect J, Majule AE. Livelihood zones analysis. A tool for planning agricultural water management investments: Tanzania. International Water Management Institute (IWMI); 2010.

30. Hammond J, Fraval S, van Etten J, Gabriel J, Mercado L, Pagella T, et al. The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. Agric Syst. 2017;151:225–233.

31. Swindale A, Bilinsky P. Household Dietary Diversity Score for measurement of household food access: Indicator guide. Food and Nutrition Technical Assistance Project, Academy for Educational Development; 2006.

32. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 4. Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K, editors. Hayama: IGES; 2006.

33. Weisell R, Dop MC. The adult male equivalent concept and its application to household consumption and expenditure studies (HCES). Food Nutr Bull. 2012;33(3):157-162.

34. Hartung C, Anokwa Y, Brunette W, Lerer A, Tseng C, Borriello G. Open Data Kit: Tools to build information services for developing regions. Proc 4th ACM/IEEE Int Conf Inf Commun Technol Dev. 2010.

35. Montpellier Panel. Sustainable intensification: A new paradigm for African agriculture. London: Agriculture for impact; 2013.

36. Smith A, Snapp S, Chikowo R, Thorne P, Bekunda M, Glover J. Measuring sustainable intensification in smallholder agroecosystems: A review. Glob Food Sec. 2017;12:127–138.

37. Snapp SS, Grabowski P, Chikowo R, Smith A, Anders E, Sirrine D, et al. M. Maize yield and profitability tradeoffs with social, human and environmental performance: Is sustainable intensification feasible? Agric Syst. 2018;162:77–88.

38. Pauw K, Thurlow J. Agricultural growth, poverty, and nutrition in Tanzania. Food Policy. 2011;36(6):795–804.

39. Frelat R, Lopez-Ridaura S, Giller KE, Herrero M, Douxchamps S, Andersson Djurfeldt A, et al. Drivers of household food availability in sub-Saharan Africa based on big data from small farms. PNAS. 2016;113(2):458–463.

40. Smith P, Martino D, Cai Z, Gwary D, Janzen H, Kumar P, et al. Greenhouse gas mitigation in agriculture. Philos Trans R Soc B. 2008;363(1492):789–813.

41. Peterman A, Behrman J, Quisumbing A. A review of empirical evidence on gender differences in nonland agricultural inputs, technology, and services in developing countries. 2010. IFPRI Discussion Paper, No. 975.

42. Tavenner K, van Wijk M, Fraval S, Hammond J, Baltenweck I, Teufel N, et al. Intensifying Inequality? Gendered trends in commercializing and diversifying smallholder farming systems in East Africa, Front Sustain Food Syst. 2019:forthcoming.

43. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974;19(6):716–723.

44. Mersmann O. emoa: Evolutionary Multiobjective Optimization Algorithms. R package version 0.5-0. 2012. Available from: https://cran.r-project.org/package=emoa

45. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.

46. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput. 2002;6(2):182–197.

47. Watson I, Marir F. Case-based reasoning: A review. Knowl Eng Rev. 1994;9(4):327–354.

48. Murdock L, Baoua I. On Purdue improved cowpea storage (PICS) technology: Background, mode of action, future prospects. J Stored Prod Res. 2014;58:3–11.

49. Megersa B, Markemann A, Angassa A, Valle Zárate A. The role of livestock diversification in ensuring household food security under a changing climate in Borana, Ethiopia. Food Secur. 2014;6(1):15–28.

50. Knight J, Wire S, Woldehanna T. The role of education in facilitating risk-taking and innovation in agriculture. J Dev Stud. 2003:39(6);1-22.

51. Berre D, Baudron F, Kassie M, Craufurd P, Lopez-Ridaura S. Different ways to cut a cake: Comparing expert-based and statistical typologies to target sustainable intensification technologies, a case-study in Southern Ethiopia. Exp Agric. 2016:online;1–17.

52. Graef F, Schneider I, Fasse A, Germer JU, Gevorgyan E, Haule F, et al. Natural resource management and crop production strategies to improve regional food systems in Tanzania. Outlook Agric. 2015;44(2):159–67.

53. Genius M, Koundouri P, Nauges C, Tzouvelekas V. Information transmission in irrigation technology adoption and diffusion: Social learning, extension services, and spatial effects. Am J Agric Econ. 2014;96(1):328–44.

54. Van Etten J, Beza E, Calderer L, Van Duijvendijk K, Fadda C, Fantahun B, et al. First experiences with a novel farmer citizen science approach: Crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). Exp Agric. 2016:online;1-22.

55. Nelson BR, Coe R, Haussmann BIG. Farmer research networks as a strategy for matching diverse options and contexts in smallholder agriculture. Exp Agric. 2016:online;1–20.

56. Sumberg J, Okali C, Reece D. Agricultural research in the face of diversity, local knowledge and the participation imperative: Theoretical considerations. Agric Syst. 2003;76:739–753.

57. Brookfield H, Parsons H, Brookfield M. Agrodiversity: Learning from farmers across the world. Canberra: Australian National University; 2003.

58. Desiere S, Jolliffe D. Land productivity and plot size: Is measurement error driving the inverse relationship? J Dev Econ. 2018;130:84–98.

59. Gladwin CH, Peterson JS, Mwale AC. The quality of science in participatory research: A case study from Eastern Zambia. World Dev. 2002;30(4):523–43.

5.3 Household-specific targeting of agricultural advice via mobile phones: Feasibility of a minimum data approach for smallholder context

Computers and Electronics in Agriculture (2019) 162:991-1000

Jonathan Steinke[abc] · Jerusha Onyango Achieng[d] · James Hammond[e] · Selamawit Sileshi Kebede[b] · Dejene Kassahun Mengistu[f] · Majuto Gaspar Mgimiloko[g] · Jemal Nurhisen Mohammed[h] · Joseph Musyoka[i] · Stefan Sieber[bc] · Jeske van de Gevel[dj] · Mark T. van Wijk[e] · Jacob van Etten[a]

[a] *Bioversity International, CGIAR Research Program on Roots, Tubers and Bananas, Turrialba, Costa Rica*

[b] *Humboldt University Berlin, Department of Agricultural Economics, Berlin, Germany*

[c] *Leibniz Centre for Agricultural Landscape Research (ZALF e.V.), Sustainable Land Use in Developing Countries, Müncheberg, Germany*

[d] *Bioversity International, Nairobi, Kenya*

[e] *International Livestock Research Institute, Nairobi, Kenya*

[f] *Bioversity International, Addis Ababa, Ethiopia*

[g] *Tanzania Agricultural Research Institute, Zonal Information and Extension Liasion Unit, Naliendele, Tanzania*

[h] *Mekelle University, Department of Dryland Crop and Horticultural Sciences, Mekelle, Ethiopia*

[i] *Lutheran World Relief, Makindu, Kenya*

[j] *University of York, York, UK*

doi.org/10.1016/j.compag.2019.05.026

Abstract

In recent years, agricultural extension services in developing countries have increasingly introduced modern information and communication technologies (ICT) to deliver advice. But to realize efficiency gains, digital applications may need to address heterogeneous information needs by targeting agricultural advisory contents in a household-specific way. We explore the feasibility of an automated advisory service that collects household data from farmers, for example through the keypads of conventional mobile phones, and uses this data to prioritize agricultural advisory messages accordingly. To reduce attrition, such a system must avoid lengthy inquiry. Therefore, our objective was to identify a viable trade-off between low data requirements and useful household-specific prioritizations of advisory messages. At three sites in Ethiopia, Kenya, and Tanzania independently, we collected experimental preference rankings from smallholder farmers for receiving information about different agricultural and livelihood practices. At each site, we identified socio-economic household variables that improved model-based predictions of individual farmers' information preferences. We used the models to predict household-specific rankings of information options based on 2-4 variables, requiring the farmer to answer between 5 and 10 questions through an ICT interface. These predicted rankings could inform household-specific prioritizations of advisory messages in a digital agro-advisory application. Household-specific "top 3" options suggested by the models were better-fit to farmers' preferences than a random selection of 3 options by 48 – 68 %, on average. The analysis shows that relatively limited data inputs from farmers, in a simple format, can be used to increase the client-orientation of ICT-mediated agricultural extension. This suggests that household-specific prioritization of agricultural advisory messages through digital two-way communication is feasible. In future digital agricultural advisory applications, collecting little data from farmers at each interaction may feed into learning algorithms that continuously improve the targeting of advice.

1. Introduction

As mobile networks and devices approach ubiquity across the Global South, agricultural extension services increasingly employ modern information and communication technologies (ICT) to deliver advice to smallholder farmers (Baumüller, 2018; ITU, 2017). Many ICT-mediated agro-information applications have recently been created around the world, such as SMS-based market information services or call centers for technical farm advice. These new services allow disseminating technical, meteorological, or market-related information to large numbers of farmers in a timely and cost-efficient manner, no matter their spatial distance to extension centers, or the advisor-farmer ratio (Aker, 2011; Baumüller, 2018; Deichmann et al., 2016). Several challenges have become apparent, however, from the implementation of the first generation of ICT-supported extension services. Disseminating generic information to farming households with heterogeneous information needs and preferences may affect the relevance and trustworthiness of advisory messages, and sometimes led to poor effects on farmers' decision-making (Aker et al., 2016; Glendenning and Ficarelli, 2012). Moreover, although delivering information through ICT is often cheaper than through conventional face-to-face extension formats, it still has a cost (Aker, 2011). Thus, to achieve desired effects on farming in a cost-efficient way, ICT applications need to specifically target disaggregated advisory contents to suitable user groups.

Through automated two-way communication interfaces, such as interactive voice response (IVR) or USSD message exchange, digital services can enable farmers to individually select preferred contents from a body of agricultural advisory messages. But the enormous variety of potential information options, especially for agronomic advice, may cause lengthy menus that can be tedious to farmers, cost time or airtime, and may thus cause attrition. Speech recognition software and artificial intelligence could help to select advisory contents according to farmers' questions, but language diversity, local dialects, and background noise cause challenges (Plauché and Nallasamy, 2007). Thus, to avoid tedious menus, while suggesting individually suitable innovation to farmers, it may be necessary to reduce the number of information options and pre-select messages that are likely to be most relevant to the user.

Agricultural extension often responds to farmers' heterogeneous information needs by targeting alternative recommendations to different types of farmers, using complex household categorizations based on characteristics such as location, resource endowments, or dominant livelihood strategy (Berre et al., 2016; Kuivanen et al., 2016). But prioritizing agricultural information for the different household categories requires extensive qualitative fieldwork, which would usually be too much effort to still warrant the efficiency gains that ICT are employed for in agricultural extension (Schindler et al., 2016). As a shortcut, information targeting can already be improved with limited, simple information about the household, such as age and gender of the household leader (Khatri-Chhetri et al., 2017). ICT applications make it possible to collect such household information remotely through users' mobile devices, and integrate the delivery of accordingly selected information in a single two-way process (Dillon, 2012; Hartung et al., 2010). It is not clear, however, how such household-specific targeting through digital channels can be done in practice. Two key decisions seem necessary: (1) which information needs to be collected from farmers, and (2) how that information should be translated into household-specific prioritizations of different agricultural advisory contents.

To achieve practical usability, an important consideration is to reduce the burden of household data collection for farmers as much as possible. But reducing the amount of household data underlying targeting may affect the fit of targeted advisory messages to households' information needs and preferences. Thus, effective use of ICT in agricultural extension implies a pragmatic balance between rapid, data-sparse household data collection and the household-specificity of advice. Effective targeting requires requesting household information from farmers that is highly predictive of their information needs as well as maximizing data quality, e.g. by recalling a low number of simple, reliable and unambiguous household indicators from farmers (Hammond et al., 2017; Jarvis et al., 2015). In this study, we investigated the feasibility of household-specific information prioritization in agricultural advisory based on simple indicators collected from farmers through ICT. Our objective was to identify a viable solution for the trade-off between minimal data enumeration and useful household-specific targeting of agricultural advisory messages.

We investigated the feasibility of such a minimum data approach to household-specific targeting in three steps. First, we used a ranking exercise to collect data on smallholder farmers' information preferences about various agricultural and livelihood development practices. We assume that a farmer's stated information preferences correspond to different expected utilities of delivering advice on these topics. Second, we fit a model to the preference data and identified household characteristics that partly explained these rankings. These characteristics were taken from a lean indicator survey, which emphasizes rapid, reliable and simple enumeration through ICT (Hammond et al., 2017). Third, we used the model to predict most likely preference rankings of further households, based on their levels of the predictor variables. These predicted preferences for information options should then inform household-specific prioritizations of advisory messages, in a two-way ICT application that collects limited data from farmers. We repeated the research process independently at three sites in Eastern Africa. By comparing the experimental stated rankings (what farmers want) and the individual predicted rankings (what the model suggests), we assessed the usefulness of our approach against an alternative scenario of no targeting. We report outcomes and discuss their implications for integrating the collection of household indicators and the prioritization of agricultural advice in a single data-sparse ICT application, such as an automated telephone line.

2. Technology background

This study on the feasibility of a minimum data approach was conceived in the context of the design of a particular digital information system. In ongoing research at three sites in Eastern Africa, we are testing a new ICT-mediated information system for sustainable intensification of smallholder agriculture. A library of audio messages about diverse agricultural topics, previously recorded by extension agents, researchers, and experienced farmers, can be accessed through telephone calls (Figure 1). To

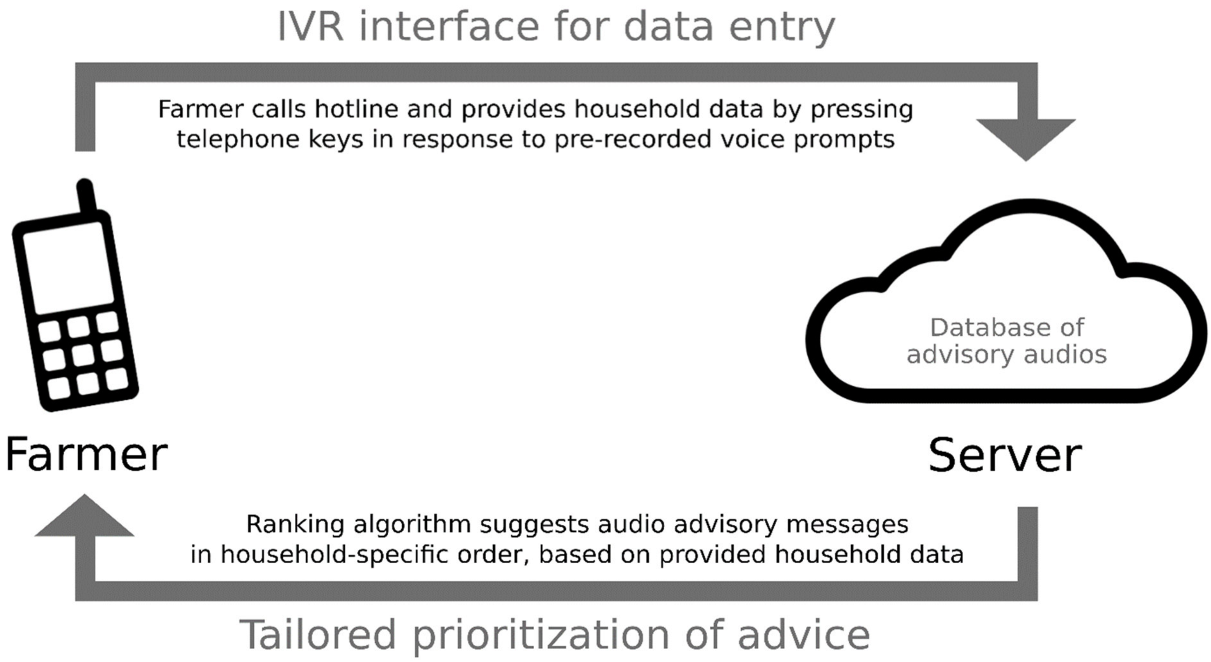

Fig. 1: Schematic overview of the intended information exchange between farmers and the online database of advisory audio messages, accessible through telephone.

decide which topics, out of a large pool of messages, to suggest to the calling farmer, the system requests the entry of household data through a hierarchic IVR menu ("Press 1 for A, press 2 for B…"). Farmers hear questions (e.g., about gender or location) and provide answers through their telephone keypads. But lengthy enumeration of household data may also cause attrition. Therefore, we were interested in minimizing the number of questions required to generate useful household-specific prioritizations of alternative advisory messages.

3. Methods

3.1 Study sites

We carried out research at three East African sites (Figure 2). By performing three independent case studies, we tested the feasibility of our approach and its robustness under contrasting circumstances. The three research sites differ in their agro-ecological and socio-economic conditions as well as in the levels of smallholder farmers' access to and experience with ICT. The Tigray region in Ethiopia is characterized by mostly arid climate and a unimodal rainfall regime, frequently experiencing droughts. About 80 % of the population depend on agriculture, which is dominated by mixed smallholder cereal-livestock systems. Food insecurity rates are high (Gebrehiwot and Van der Veen, 2013). Makueni County in Kenya has predominantly semi-arid climate and a bimodal rainfall pattern, with recurrent drought events. Farming systems are primarily based on maize, cow pea, green grams, and grazing livestock (Speranza et al., 2010). The Southern Agricultural Zone in South-Eastern Tanzania comprises the administrative regions of Lindi and Mtwara, as well as Tunduru

District of Ruvuma Region. Climate is tropical with a varyingly uni/bimodal rainfall distribution. Agriculture concentrates on maize, cassava, and pulses for subsistence and commercial production of oil seeds (Perfect and Majule, 2010). Yields of staple crops are among the lowest at country level (Rowhani et al., 2011). In the remainder of this study, the sites are referred to by the country they are situated in.

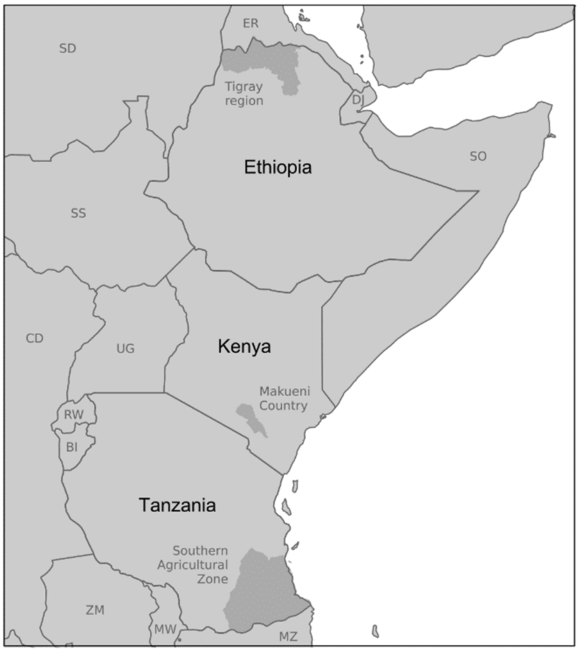

Fig. 2: Research sites in Eastern Africa. Neighboring countries are marked with ISO two-letter country codes. Spatial data retrieved from gadm.org.

3.2 Household surveys

Because we were interested in linking farmers' preferences for receiving different advisory contents with household characteristics, we first carried out country-specific variants of the "RHOMIS" lean indicator household survey (Hammond et al., 2017). This survey was designed for ICT-mediated enumeration using Open Data Kit software (Hartung et al., 2010), and intends to minimize respondent fatigue and resulting data inaccuracy by using simple questions about observable criteria. The data included variables related to household composition, resources, and the farming system. At each site, enumerator teams used smartphones to collect the data. Households were randomly sampled from beneficiary villages involved in an ongoing research project led by Bioversity International by sampling a country-wise constant number of smallholder farmers per village. 249 households were successfully surveyed in Ethiopia, 316 households in Kenya, and 521 households in Tanzania. Median farmer-stated land holdings were 0.61 ha in Ethiopia, 2.43 ha in Kenya, and 2.84 ha in Tanzania.

3.3 Experimental elicitation of farmers' information preferences

To determine farmers' individual information preferences at each site, we used a choice experiment. Farmers were asked to rank 9 different household-level practices according to their interest in receiving more information about them. We then used these stated preferences to train a recommendation system.

As information options in the choice experiments, we prepared sets of practices that were locally viable but not yet widely adopted by farmers in the area. These selections included innovative or rare practices found with so-called "positive deviant" households (Steinke et al., 2019). The fact that these strategies have before been implemented by relatively successful farmers makes them likely to be generally interesting options for further farmers, although not all options may appear equally

Fig. 3: Enumeration of farmers' information preferences in Ethiopia

suitable to all farmers. Simpler methods could also be used to produce a list of information options, such as quick elicitations from lead farmers, experienced extension agents, or agricultural researchers. In the context of this study, however, our approach ensured that, for each site, there was a set of information options with a similar level of local relevance. The procedure we followed to identify the practices is described in more detail in the supplementary information to this article.

Through a simple ranking experiment, we then determined farmers' individual preferences for information about 9 alternative information options. All options were illustrated on individual, roughly hand-sized cards. We randomly sampled household leaders from the initial RHOMIS survey to become participants in our ranking experiments (n = 86 in Ethiopia, n = 43 in Kenya, n = 98 in Tanzania). We asked participants to order the cards in accordance to how strongly they would like to learn more about the illustrated practices and recorded the ranking orders (Figure 3). In most cases, this involved further on-spot explanations about the practices by the enumerators. For data exploration, we analyzed the internal heterogeneity of rankings at each study site by Kendall's *W*, a coefficient of rank concordance (Kendall and Babington Smith, 1939), using the package irr (Gamer et al., 2012) in the R software (R Core Team, 2018). We interpreted Kendall's *W* using the classification system by Schmidt (1997).

3.3 Analysis of preference data

3.4.1 Estimation of overall most likely rankings of information options

At each site, we first identified the most likely overall preference ranking across all respondents (n = {86, 43, 98}) by fitting a Bradley-Terry model to farmers' stated rankings (Bradley and Terry, 1952). Bradley-Terry models identify the overall most likely order from multiple rankings of the same items. Because Bradley-Terry models rely on pairwise comparison data, we first converted the rankings to a pairwise comparison data format. Converting rankings to pairwise comparisons involves an information loss, but allows statistical analysis with covariates (ranker characteristics), using the generalized linear model framework (Dittrich et al., 2000). In contrast to the Bradley-Terry model, the Plackett-Luce model analyzes rankings directly (Luce, 1959; Plackett, 1975). Currently available implementations of the Plackett-Luce model, however, do not follow the generalized linear model framework and the partitioning-based framework has limited statistical power (Turner et al. 2018). To get a quantitative idea of the potential information loss caused by converting rankings to pairwise comparisons, we compared rankings and preference scores generated by Bradley-Terry models and Plackett-Luce models, respectively (for detail, see following Section). We used the packages BradleyTerry2 (Turner and Firth, 2012) and PlackettLuce (Turner et al., 2018) in the R software (R Core Team, 2018). The maximum likelihood parameter estimates (log-odds) of the practices ranked by each Bradley-Terry and Plackett-Luce models had Pearson's correlation coefficients between 0.77 (Tanzania) and 0.96 (Ethiopia), suggesting that the information loss is moderate to small.

3.4.2 Estimation of overall preference scores of information options

The Bradley-Terry model uses maximum likelihood to estimate the log-odds of options being

Table 1

Candidate covariates used in specification of Bradley-Terry models of farmers' information preferences

Variable category	Variable	Definition (unit)	Number of survey questions needed
Basic household variables	Gender of household head	Female, Male	1
	Age of household head	(years)	1
	Region	2 options in Ethiopia, 1 in Kenya, 2 in Tanzania	1
Resources	Land holdings	(ha)	1
	Labour availability	Household size (in MAE) divided by land holdings	7
	Livestock holdings	(Tropical livestock units)	6
	Social capital	First loading of a principal component analysis on indicators of membership in established groups, and access to public benefits	3
Farming style-related	Land tenure	Household owns land: yes/no	1
	Labour hiring	Household ever hires workers for farming: yes/no	1
	Input changes	Household has changed the use of agricultural inputs over the last year: Decrease/No change/Increase	1

ranked higher than a reference option, which is arbitrarily set to 0. We converted these values into probabilities, and then calculated, for each information option, the probability of being ranked higher than all other options (the relative "preference score") by iteratively modifying the reference, following the procedure described by Jeske et al. (2007). We then identified sets of practices that were ranked significantly different by the farmers by testing which of the pairwise differences in preference scores of practices were significantly different from 0. For this, we corrected the p-values for multiple comparisons using Holm's sequential Bonferroni procedure (Holm, 1979).

3.4.3 Model specification with household variables

Our ultimate goal was to predict the most likely individual preference rankings for further target households. These predicted rankings would then inform household-specific prioritizations of advisory messages. For this, we needed models that linked rankings with household characteristics. Therefore, we further specified the Bradley-Terry models by introducing socio-economic household variables as covariates. Candidate covariates were selected following two criteria. Our first criterion was that variables should be known to affect the applicability of specific agricultural practices and/or farmers' preferences for agricultural information (e.g., Berre et al., 2016; Kassie et al., 2016; Khatri-Chhetri et al., 2017). Our second criterion was that the variables should be based in a limited number of simple questions, to allow rapid data collection through a digital interface. We did not consider variables that require more than 7 separate question in the RHOMIS framework (see Section 3.2 above). This criterion meant we did not consider some potentially important variables, such as financial resources or market orientation, for which more detailed series of questions are required to generate reliable data (Hammond et al., 2017; Hanisch, 2005). The resulting selection of candidate covariates is shown in Table 1. These included three basic household variables (gender, age, region), four proxies of productive resource availability, and three variables reflecting (dis-)investments into agricultural intensification, roughly corresponding to different "farming styles" (Van der Ploeg and Ventura, 2014). For Ethiopia and Tanzania, there were 10 candidate variables, while for Kenya there were 9. In Kenya, the survey covered only one administrative region, so region was omitted as a covariate for Kenya.

We then specified models by forward variable selection using the "Permuted Inclusion Criterion" (Lysen, 2009). This procedure consists of two steps. In the first step, we added to the set of original covariates an additional set of fake variables generated by randomly permuting the original variables. As a result, every farmer ranking of practices was linked to a set of observed variables and a set of permuted variables, i.e. the characteristics of another randomly selected farmer. Permuted variables were not expected to have any predictive power for rankings. In the second step, we added covariates to the Bradley-Terry model. We added each variable (real and permuted) to the null model separately and recorded which of the variables reduced model deviance most strongly. We replicated this process 500 times, each time with a new random permutation. Across the 500 runs, we identified the covariate that appeared most often as the most deviance-reducing one. When this was

Table 2

Agricultural and livelihood practices identified with "positive deviant" households and mean Bradley-Terry parameter estimates for farmers' preference rankings of information about these practices. In groupings of practices, different letters indicate significantly different ranks of information options.

Information option [a]	Code (Figure 4)	Kendall's *W* of all rankings	Preference score	Grouping
Ethiopia (n = 86)		0.482		
Sowing cereals in lines	L		0.806	a
Diligent farm scheduling	S		0.780	a
Rain water harvesting	R		0.671	b
Storing and trading crops	T		0.512	c
Opening a business	B		0.375	d
Tree nursery	N		0.361	d
Reducing food wastage	W		0.351	d
Finding off-farm job	J		0.329	d
Improving crop storage	C		0.314	d
Kenya (n = 43)		0.495		
Machine tillage	M		0.764	a
Terracing	T		0.726	a
Zai pits	Z		0.712	a
Dry planting	D		0.673	a
Collective crop marketing	G		0.500	b
Mulching	R		0.438	b
Opening a business	B		0.380	b
Renting out traction animals	O		0.168	c
Finding off-farm job	J		0.139	c
Tanzania (n = 98)		0.318		
Intercropping Pigeon pea / Maize	I		0.675	a
Improving crop storage	C		0.645	a
Diligent farm scheduling	S		0.636	a
Machine tillage	M		0.492	b
Intensifying poultry production	P		0.460	b
Opening a business	B		0.450	b
Tree nursery	T		0.427	b
"Livestock bank"	L		0.426	b
Finding off-farm job	J		0.287	c

[a] For explanations about the practices see the supplementary information to this article

a real variable, we added it to the model, excluded the corresponding permuted variable from data, and continued forward selection. We stopped co-variate selection when a permuted variable was found to be the most frequent most deviance-reducing variable, i.e., when no real variable had more explanatory power than the fake ones. The relative influence of different household characteristics on farmers' preferences was quantified by the respective step-wise changes in model deviance caused by including each variable in the model. We re-scaled the values by setting the highest value to 1.

We assessed goodness-of-fit of the models by reduction in model deviance compared to the null model (no covariates). In addition, we calculated the mean pairwise agreement between individual stated rankings and the rankings predicted for the same farmers based on their household characteristics. For this, we used Kendall's tau, a coefficient of similarity between two rankings (Kendall, 1938). Kendall's tau can take values from -1 (inverse ranking) to +1 (identical ranking). We used the package Kendall (McLeod, 2011) in the R software (R Core Team, 2018).

3.5 Generating household-specific prioritizations of information options

As a final step, we used the fit models to predict the most likely preference rankings for all households enumerated in the RHOMIS surveys (n = {249, 316, 521}, see Section 3.2). This generated a household-specific prioritization of the information options for each household, based on the characteristics previously identified as predictors.

We assessed the usefulness of these household-specific prioritizations in three ways, always comparing farmers' stated preference rankings (training data from n = {43, 86, 98} farmers) and the household-specific model outputs for these same farmers. Firstly, we calculated the mean Kendall's rank correlation (Kendall's tau) between stated and predicted preference rankings (see above). Secondly, we specifically explored the consequences of using the prioritizations to make individual "top 3" suggestions to target households. We assessed the match between the 3 options ranked highest by respondents, and the "top 3" suggested by the fit models for these particular farmers by counting the options in agreement, regardless of the particular rank positions within each set of three. Thirdly, we differentiated these agreement scores by the 9 information options. For each option, we calculated the probability of being correctly included in the "top 3" suggestions for respondents who had included that practice in their "top 3" preferences.

To compare the model-based targeting approach with a no-targeting alternative, we also assessed the usefulness of random prioritizations. For this, we generated a random order of the information options for each household and performed the same three steps of analysis as for the model-based prioritizations. We repeated this process 1000 times and always calculated mean scores from 1000 runs.

4. Results

At all study sites, farmers expressed heterogeneous preferences for agricultural information (Figure 4, left side). There was moderate overall agreement in ranking the information options among Ethiopian and Kenyan respondents (Kendall's $W \approx 0.5$), but preferences were more differentiated in Tanzania (Table 2). Nonetheless, at all sites, Bradley-Terry

Table 3

Goodness-of-fit parameters of Bradley-Terry models of farmers' information preferences. Predictor weights represent relative reductions in residual deviance through a deviance-based forward selection procedure and are scaled by setting the maximum value to 1.

Model parameters	**Ethiopia**	**Kenya**	**Tanzania**
Null deviance	3693.1	1297.6	4541.5
Residual deviance	2858.0	845.5	4144.0
Degrees of freedom	2616	904	3236
Mean Kendall's tau between stated and predicted rankings	0.47	0.38	0.30
Predictor weights			
Age	0.728		0.443
Administrative region	0.615		
Labour availability	0.821	0.956	
Social capital			0.104
Input changes	1.000	1.000	1.000

models identified significantly different preference scores for the information options (Table 2). In Ethiopia, practices could be categorized into four distinct groups with significant differences between their positions in farmers' rankings. In both Kenya and Tanzania, there were three groups of practices (Table 2).

At each site, farmers' rankings were associated with certain socio-economic characteristics (Table 3). A specific set of two to four household characteristics reduced Bradley-Terry model deviance and explained part of the variation in preferences for agricultural information. Variables that partly explained preferences included: Age of the household head, Region, Labour availability, Social capital, and a recent change in agricultural input use. Of the 10 variables we tested, however, 5 did not contribute to model fit in any of the country cases: Gender, Land holdings, Livestock holdings, Land tenure, and Labour hiring.

Using the identified household variables as predictors, the Bradley-Terry models determined a most likely preference ranking for each surveyed household (Figure 4, right side). These predicted rankings were less differentiated than the stated rankings, with Kendall's W of 0.85 in Ethiopia, 0.86 in Kenya, and 0.81 in Tanzania. On average, pairwise agreement between farmers' stated preference rankings and model-predicted rankings based on the respective farmer's characteristics was moderate to strong (mean Kendall's tau ranging from 0.30 to 0.47, Table 3). These predicted household-specific prioritizations varied according to the households' characteristics: For example, for Ethiopian households that had recently *increased* their agricultural input use, predictions set the option "Finding an off-farm job" at an average rank of 7.7. For households that had recently decreased input use, this option was deemed more suitable, with an average predicted rank of 4.3. In Tanzania, the Bradley-Terry model suggested "Intercropping maize/pigeon pea" as top option for 83 % of the recent input *increasing* households, whereas it gave highest priority to "Improving crop storage" for all input *decreasing* households.

Comparing the stated rankings with both random rankings and model-predicted rankings showed that household-specific "top 3" information options suggested by the models were better fit to farmers' preferences than the "top 3" of a random order (Table 4). Suggesting to each farmer a random selection of 3 out of 9 options would include, on average, 1 of the farmer's three most-preferred options. With household-specific prioritizations generated by the fit Bradley-Terry models, the "top 3" options included an average of 1.48 (Tanzania) to 1.68 (Kenya) of the farmers' three most-preferred options (regardless of the specific rank within the set of three). Across all tested households at all three sites, this mean agreement between stated

Table 4

Selecting "top 3" suggested information options either by Bradley-Terry models or at random: Mean agreement with farmer-ranked top 3, and probabilities of individual information options being correctly included in "top 3" suggestions.

		Ethiopia			Kenya			Tanzania	
		n=82[a]			n=31[a]			n=91[a]	
		Targeting	Random		Targeting	Random		Targeting	Random
Mean agreement between observed and predicted preferences		1.55	1.00		1.68	1.00		1.48	1.00
Number of practices correctly included in "top 3"	3	9%	1%		13%	7%		5%	1%
	2	43%	21%		52%	42%		44%	22%
	1	44%	53%		26%	40%		44%	54%
	0	5%	24%		10%	11%		7%	24%
Information option suggested adequately	L	22%	30%	M	100%	20%	I	93%	32%
	S	74%	30%	T	67%	24%	C	84%	31%
	R	67%	31%	Z	69%	25%	S	97%	29%
	T	78%	31%	D	74%	20%	M	0%	33%
	B[b]	-	-	G	0%	26%	P	15%	32%
	N	30%	31%	R	0%	19%	B	0%	34%
	W	57%	32%	B	0%	18%	T	0%	29%
	J	0%	31%	O	0%	34%	L	0%	31%
	C	0%	34%	J	0%	11%	J	0%	30%

[a] Numbers of predictions are lower than numbers of recorded observations (Table 1) due to missing household data for some ranking households

[b] No ranking household had included this option in their top 3

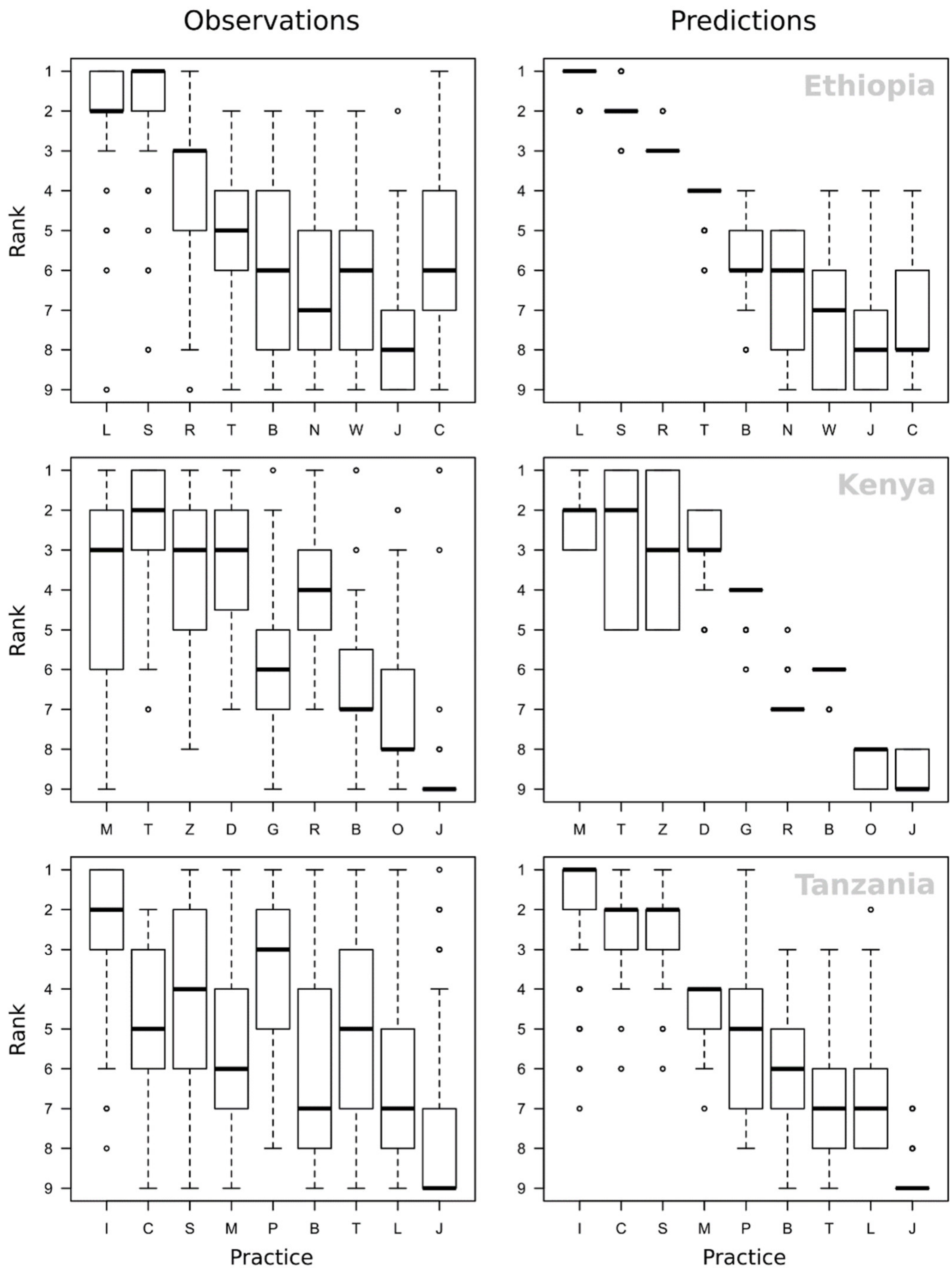

Fig. 4: Stated rankings (left) and rankings predicted by the fit Bradley-Terry models (right). For the practice codes on horizontal axes, see Table 2. n(observations) = 86 in Ethiopia, 43 in Kenya, and 98 in Tanzania. n(predictions) = 249 in Ethiopia, 316 in Kenya, and 521 in Tanzania.

and model-predicted "top 3" options was 1.54. With model-based targeting, the probability of suggesting to farmers at least 2 out of their 3 most-preferred options was more than doubled in

Ethiopia (a 52 % chance instead of 22 % without targeting) and Tanzania (49 % instead of 23 %). In Kenya, where farmers' preferences showed stronger variation among the most-preferred information options, the relative benefit of model-based targeting over random suggestions was weaker, but still evident (65 % versus 49 % without targeting). At all sites, targeting reduced the probability of a "complete miss", i.e. including none of the farmers' 3 most-preferred options in the "top 3" suggestion. In Ethiopia, for example, the probability for this to happen was 5 %, compared to 24 % in a no-targeting scenario.

5. Discussion

5.1 Small sets of household variables help to predict information preferences

This study demonstrates that relatively little household data can be sufficient to anticipate farmers' individual preferences for agricultural information in a way that allows usefully customized prioritizations of advisory messages. Although predicted rankings were not perfectly congruent with observed preferences, the models made household-specific suggestions that were, on the whole, better-fit to farmers' preferences than random recommendations. The socio-economic household variables associated with information preferences differed between sites, which also involved different tested portfolios of information options. But ooverall, having implemented a recent change in agricultural input use, such as chemical fertilizer or improved seeds, was the strongest predictor across all sites, as well as the only universal one. This suggests that a household's "farming style" may be more important information for prioritizing household-specific development strategies than its access to productive resources, which many farm typologies rely on. Indeed, despite similar resource endowments, farmers may seek highly diverse development strategies, e.g. in function of their risk aversion or the dominant output sought after, such as increasing cash income or sustaining food production (Van der Ploeg and Ventura, 2014). This finding has implications for the design of digital extension applications that target advice: Enumerating household resource endowments through ICT may be easier than collecting information on farming styles, which can be hard to collect through numeric data or yes/no questions (Fairweather and Klonsky, 2009). Nevertheless, our analysis suggests that adequate targeting of advice should use data on target farmers' farming styles. This could include, for example, information about fertilizer purchases or recent on-farm investments.

5.2 Useful prioritization of advisory messages based on data enumerated through ICT seems feasible

To assess the usefulness of the model-based targeting approach presented here, an important question is whether it can reduce the risk of disseminating information of low relevance. This is a crucial criterion for the design of digital advisory services (Nakasone et al., 2014). Our analysis explored the scenario of delivering customized "top 3" suggestions of agricultural advisory contents. Compared to random suggestions, the share of farmers receiving predominantly irrelevant messages was greatly reduced at each site (e.g., from 51 % down to 36 %

in the weakest case, Kenya). Overall, through the targeting approach, a majority of households received “top 3” suggestions that were better-fit to their preferences than random orders.

Although an initial data collection effort is needed to train the first model, the benefit of delivering targeted advice to a large number of households may justify the execution of the ranking exercise with a limited number of farmers. Because predictor variables are not universal, model predictions are valid only for the study region, and only for the practices originally included. In the future, analysis may be refined by fitting local sub-models through recursive partitioning (Strobl et al., 2011). Moreover, linking preferences to objective characteristics of practices (e.g., implementation costs, expected effects on labour availability) may allow introducing new practices to the prioritization model and a resulting digital information service, without repeating the ranking experiment. This study demonstrates that even with a relatively small sample of farmers training the initial model, improved targeting of a set of initial advisory messages is possible. Over time, as farmers start using an ICT-mediated information system and make choices – e.g. about the most-preferred out of a set of three promoted practices – the household-specific suggestions of promoted practices could be further refined. Each time a farmer calls, they might be asked 1-2 additional questions about their household and farming system. As the sample size grows and more household data, as well as partial ranking choices, enter the model, the system will increase its predictive power, potentially also using more predictor variables not included in this study. An initial targeting model, informed by the choice experiment with representative households, would be needed to offer first-time users an acceptable experience, to encourage usage of the service. Over time, learning algorithms or regular manual adjustments to the model should use newly accumulating data to continue to improve the targeting of agricultural advice.

But does the improvement in targeting advice justify the enumeration effort on the farmer side? At each site, the models generated prioritizations based on two to four household variables. These variables were calculated from sets of 5 (Tanzania) to 10 (Ethiopia) questions. The most important variable, recent changes to agricultural input use, requires only one question. Mini-questionnaires of a few questions can be implemented through ICT, e.g. via USSD menus or interactive voice response, both of which request users to enter data through the keypad of conventional mobile phones (“Press 1 for topic A, press 2 for topic B …”). Through recent developments in mobile money services, mobile phone users across the Global South are becoming increasingly acquainted with these technologies (GSMA, 2017). Designers of new agro-advisory services will need to identify a viable trade-off between questionnaire length and predictive power of the information for household-specific targeting of advisory contents. Our results suggest that prioritization of advice through ICT tools is possible, and that a satisfactory trade-off can be achieved between rapid, simple household data enumeration and useful household-specific prioritizations. The rise of smartphone ownership among rural population worldwide likely offers even more opportunities for household- and even plot-specific targeting of agricultural advice, taking additional benefit of features such as GPS or video (Carmona et al., 2018).

Household data used in this study was collected using ICT (Open Data Kit on mobile Android devices), but not entered by farmers themselves. Although the lean indicators in the RHOMIS survey were designed for simple and unambiguous enumeration, this might mean that farmers can face unexpected difficulties in providing the requested household information without prior training (Lerer et al., 2010; Patnaik et al., 2009). In ongoing research, we are observing farmers' interaction with the IVR interface, in order to make necessary adaptations to the sequence of data entry or IVR voice prompts.

5.3 Farmers' overall information preferences can inform priority-setting for advisory services

Our results suggest that information on farmers' information preferences, which may also accumulate as farmers use a digital agro-advisory application and make choices, can generate more general, useful insights for advisory services. Despite heterogeneity in respondents' rankings, at each site, the Bradley-Terry models identified distinct groups of practices that were given significantly different priority by the farmers. Such categorization of information options by overall popularity can be useful for extension services, e.g. to select topics about which to provide particularly detailed information. For example, strong overall interest in line sowing in Ethiopia may warrant providing multiple, crop-specific messages about line spacing. Because there is a trade-off between the need for disaggregating information according to farmers' preferences and the costs of generating contents, knowing which topics to emphasize in greater detail can be important for the financial sustainability of digital advisory applications (Nakasone et al., 2014).

Across all sites, practices related to own agricultural production were generally preferred over non-agricultural options. This finding underlines the need for advisory support to established household activities, rather than diversification of rural livelihoods. In particular, "Finding off-farm job" was of little interest to the responding farmers. This seems to contrast calls for supporting non-agricultural income options in rural development, which are often based on sound econometric analysis (e.g. Frelat et al., 2016), but may face challenges in practice due to farmers' livelihood preferences and aspirations (Verkaart et al., 2018). Our results support the idea that pure information interventions without practical demonstration activities – such as the provision of audio messages through a hotline – may be most effective by focusing on knowledge-based, gradual modifications of current systems. When farmers need to make investments, e.g. in labour or machinery, information interventions may nevertheless need to be accompanied by additional measures, such as insurance schemes (Pradhan et al., 2015).

6. Conclusions

This study demonstrates the feasibility of useful household-specific prioritizations of agricultural information based on small sets of household indicators collected through ICT. Although training the first models with experimental and survey data from representative households requires an initial effort, this may contribute to resource-efficient strategies of engaging ICT in agricultural extension. We found that it is possible to achieve a satisfactory trade-off between minimal data enumeration, which is required if farmers are to use ICT for access to advisory services, and the household-specific adaptation of advice. This approach is especially useful to deliver a first set of relevant content to farmers, who could be asked for some information when registering to the service. Once farmers start using the service, the digital system itself may continuously generate new data about users' preferences and characteristics, thus improving the model-based targeting with new training information.

In the context of the particular digital solution we are considering (Section 2), this supports the idea that it is feasible to deliver individually targeted agricultural information to heterogeneous households through an automated call-in hotline connected to a database of audio records. An interactive voice response menu, requesting farmers to answer a low number of questions using their telephone's keypad, may enable ICT applications of this kind to select suitable advisory contents. To justify investments into new services, further research needs to establish to what extent a household-tailored advisory application increases adoption and continued use of promoted practices, compared to more "one-size-fits-all" approaches to agricultural advisory. Our results are also relevant for other applications that involve household-specific agricultural advice. In the future, research may produce more generalizable insights about which data-sparse indicators can serve as predictors of farmers' information needs. Small standard sets of questions that efficiently capture the factors behind farmers' information needs will likely be useful for a wide range of digital applications in agricultural advisory.

Acknowledgements

This research was undertaken as part of, and funded by, the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Trust Fund contributors, see https://www.cgiar.org/funders/. Funding support for this work was also provided by UK Aid from the UK government through the Sustainable Agricultural Intensification Research and Learning in Africa programme (SAIRLA). However, the views expressed do not necessarily reflect the UK government's official policies. Additional funding was kindly provided by Stiftung fiat panis. We owe sincere gratitude to all farmers and the many field enumerators contributing to this study.

References

Aker, J.C., 2011. Dial "A" for agriculture: A review of information and communication technologies for agricultural extension in developing countries. Agric. Econ. 42, 631–647.

Aker, J.C., Ghosh, I., Burrell, J., 2016. The promise (and pitfalls) of ICT for agriculture initiatives. Agric. Econ. 47, 35-48.

Baumüller, H., 2018. The little we know: An exploratory literature review on the utility of mobile phone-enabled services for smallholder farmers. J. Int. Dev. 154, 134–154.

Berre, D., Baudron, F., Kassie, M., Craufurd, P., Lopez-Ridaura, S., 2016. Different ways to cut a cake: Comparing expert-based and statistical typologies to target sustainable intensification technologies, a case-study in Southern Ethiopia. Exp. Agric. (First View), 1–17.

Bradley, R.A., Terry, M.E., 1952. Rank analysis of incomplete block designs. I. The method of paired comparisons. Biometrika 39, 324–345.

Carmona, M.A., Sautua, F.J., Pérez-Hernández, O., Mandolesi, J.I., 2018. AgroDecisor EFC: First Android™ app decision support tool for timing fungicide applications for management of late-season soybean diseases. Comput. Electron. Agric. 144, 310–313.

Deichmann, U., Goyal, A., Mishra, D., 2016. Will digital technologies transform agriculture in developing countries? Agric. Econ. 47, 21–33.

Dillon, B., 2012. Using mobile phones to collect panel data in developing countries. J. Int. Dev. 24, 518–527.

Dittrich, R., Katzenbeisser, W., Reisinger, H., 2000. The analysis of rank ordered preference data based on Bradley-Terry type models. OR Spektrum 22, 117–134.

Fairweather, J.R., Klonsky, K., 2009. Response to Vanclay et al. on farming styles: Q Methodology for identifying styles and its relevance to extension. Sociol. Ruralis 49, 189–198.

Frelat, R., Lopez-Ridaura, S., Giller, K.E., Herrero, M., Douxchamps, S., Andersson Djurfeldt, A., Erenstein, O., Henderson, B., Kassie, M., Paul, B.K., Rigolot, C., Ritzema, R.S., Rodriguez, D., van Asten, P.J.A., van Wijk, M.T., 2016. Drivers of household food availability in sub-Saharan Africa based on big data from small farms. PNAS 113, 458–463.

Gamer, M., Lemon, J., Fellows, I., Singh, P., 2012. irr: Various coefficients of interrater reliability and agreement. R package version 0.84.

Gebrehiwot, T., van der Veen, A., 2013. Climate change vulnerability in Ethiopia: Disaggregation of Tigray Region. J. East. African Stud. 7, 607–629.

Glendenning, C.J., Ficarelli, P.P., 2012. The relevance of content in ICT initiatives in Indian agriculture. IFPRI Discussion Papers (No. 01180), Washington, D.C.

GSMA, 2017. State of the Industry Report on Mobile Money. GSM Association, London.

Hammond, J., Fraval, S., van Etten, J., Gabriel, J., Mercado, L., Pagella, T., Frelat, R., Lannerstad, M., Douxchamps, S., Teufel, N., Valbuena, D., van Wijk, M.T., 2017. The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. Agric. Syst. 151, 225–233.

Hanisch, J.U., 2005. Rounded responses to income questions. All. Stat. Arch. 89, 39–48.

Hartung, C., Anokwa, Y., Brunette, W., Lerer, A., Tseng, C., Borriello, G., 2010. Open Data Kit: Tools to build information services for developing regions. Proc. 4th ACM/IEEE Int. Conf. Inf. Commun. Technol. Dev.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, 65–70.

ITU, 2017. ICT facts and figures 2017, International Telecommunications Unit.

Jarvis, A., Eitzinger, A., Koningstein, M.J., Banjamin, T., Howland, F.C., Andrieu, N. V, Twyman, J., Corner-Dolloff, C., 2015. Less is more: The 5Q approach. CIAT, Cali, Colombia.

Jeske, D.R., Lesch, S.M., Deng, H., 2007. The merging of statistics education, consulting and research: A Case Study. J. Stat. Educ. 15, 1–19.

Kassie, M., Zikhali, P., Manjur, K., Edwards, S., 2009. Adoption of sustainable agriculture practices: Evidence from a semi-arid region of Ethiopia. Nat. Resour. Forum 33, 189–198.

Kendall, M.G., 1938. A new measure of rank correlation. Biometrika 30, 81–93.

Kendall, M.G., Babington Smith, B., 1939. The problem of m rankings. Ann. Math. Stat. 10, 275–287.

Khatri-Chhetri, A., Aggarwal, P.K., Joshi, P.K., Vyas, S., 2017. Farmers' prioritization of climate-smart agriculture (CSA) technologies. Agric. Syst. 151, 184–191.

Kuivanen, K.S., Alvarez, S., Michalscheck, M., Adjei-Nsiah, S., Descheemaeker, K., Mellon-Bedi, S., Groot, J.C.J., 2016. Characterising the diversity of smallholder farming systems and their constraints and opportunities for innovation: A case study from the Northern Region, Ghana. NJAS – Wageningen J. Life Sci. 78, 153–166.

Lerer, A., Ward, M., Amarasinghe, S., 2010. Evaluation of IVR data collection UIs for untrained rural users. Proc. First ACM Symp. Comput. Dev. - ACM DEV '10.

Luce, R.D., 1959. Individual choice behavior: A theoretical analysis. Wiley, New York.

Lysen, S., 2009. Permuted inclusion criterion: A variable selection technique. University of Pennsylvania.

McLeod, A.I., 2011. Kendall: Kendall rank correlation and Mann-Kendall trend test. R package version 2.2.

Nakasone, E., Torero, M., Minten, B., 2014. The power of information: The ICT revolution in agricultural development. Annu. Rev. Resour. Econ. 6, 533–550.

Patnaik, S., Brunskill, E., Thies, W., 2009. Evaluating the accuracy of data collection on mobile phones: A study of forms, SMS, and voice. Proc. Int. Conf. Inf. Commun. Technol. Dev. (ICTD).

Perfect, J., Majule, A.E., 2010. Livelihood zones analysis. A tool for planning agricultural water management investments: Tanzania. International Water Management Institute (IWMI).

Plackett, R., 1975. The analysis of permutations. J. R. Stat. Soc. Ser. C (Applied Stat.) 24, 193–202.

Plauché, M., Nallasamy, U., 2007. Speech interfaces for equitable access to information technology. Inf. Technol. Int. Dev. 4, 69-86.

Pradhan, P., Fischer, G., van Velthuizen, H., Reusser, D.E., Kropp, J.P., 2015. Closing yield gaps: How sustainable can we be? PLoS One 10, e0129487.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rowhani, P., Lobell, D.B., Linderman, M., Ramankutty, N., 2011. Climate variability and crop production in Tanzania. Agric. For. Meteorol. 151, 449–460.

Schindler, J., Graef, F., König, H.J., Mchau, D., Saidia, P., Sieber, S., 2016. Sustainability impact assessment to improve food security of smallholders in Tanzania. Environ. Impact Assess. Rev. 60, 52–63.

Schmidt, R.C., 1997. Managing delphi surveys using nonparametric statistical techniques. Decis. Sci. 28, 763–774.

Speranza, C.I., Kiteme, B., Ambenje, P., Wiesmann, U., Makali, S., 2010. Indigenous knowledge related to climate variability and change: Insights from droughts in semi-arid areas of former Makueni District, Kenya. Clim. Change 100, 295–315.

Steinke, J., Mgimiloko, G., Graef, F., Hammond, J., van Wijk, M.T., van Etten, J. Prioritizing options for multi-objective agricultural development through the Positive Deviance approach. PLoS One 14, e0212926.

Strobl, C., Wickelmaier, F., Zeileis, A., 2011. Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. J. Educ. Behav. Stat. 36, 135–153.

Turner, H., Kosmidis, I., Firth, D., van Etten, J., 2018. PlackettLuce: Plackett-Luce models for rankings. R package version 0.2-2.

Turner, H.L., Firth, D., 2012. Bradley-Terry models in R: The BradleyTerry2 package. J. Stat. Softw. 48, 1-21

Van der Ploeg, J.D., Ventura, F., 2014. Heterogeneity reconsidered. Curr. Opin. Environ. Sustain. 8, 23–28.

Verkaart, S., Mausch, K., Harris, D., 2018. Who are those people we call farmers? Rural Kenyan aspirations and realities. Dev. Pract. 28, 468–479.

6 Conclusion

6.1 Discussion of results

This section returns to the research questions (Section 3) to discuss the empirical evidence presented in the research publications. It critically assesses the relevance of the generated insights for consolidating viable methodologies based on the three solution concepts proposed here: Agricultural citizen science, Positive Deviance, and digital two-way communication. Practical implications for integrating this new knowledge into methodological development within agricultural extension services are discussed in the subsequent section.

Research question 1

How can ICT-mediated agricultural citizen science help to involve large numbers of smallholder farmers in knowledge generation?

Over the last two decades, the environmental and biomedical sciences have established a variety of methodologies for scaling public participation in research, often relying on digital interactions through online platforms. Citizen science approaches for agricultural research, especially for smallholder context in developing countries, have been explored just recently. As such, the 'tricot' methodology for crowdsourcing variety selection was proposed by Van Etten (2011) as a scalable approach to engage potentially large numbers of farmers in on-farm technology testing. In this methodology, farmers test sets of three technologies, such as seed varieties, and record comparative observations (thus *triadic comparisons of technologies*, tricot).

The results presented in this work (Steinke et al. 2017) indicate that the tricot methodology can generate valid scientific insights on the relative qualities of multiple varieties of common bean (*Phaseolus vulgaris* L.). Such findings can be useful both for local innovation processes, such as identifying locally suitable varieties, as well as for wider scientific progress, such as identifying promising germplasm for plant breeding efforts (van Etten et al. 2019a). Agricultural scientists may thus be encouraged to consider tricot as a research tool by the opportunities associated with this multi-location trial format, such as rapidly replicating technology testing across gradients of climate or management.

For Ethiopia, India, and Honduras, Beza et al. (2017) show that smallholder farmers participating in tricot trials are motivated to contribute to knowledge generation by a variety of motives. For example, farmers are incentivized by the sense of helping the researchers, growing their social network, and contributing to community development. Thus, creating these incentives for farmers seems crucial to successful tricot-style research. The apparent discordance between the individual nature of on-farm trials and the highly social nature of farmer-perceived benefits from participation seems surprising at first. This suggests that agricultural citizen science, like other forms of farmer participation in research, needs to be embedded in social organizations, such as community-based farmer groups or extension services.

Due to the 'Wisdom of Crowds' principle demonstrated in this work, tricot is not only a methodology to enable scaling farmer participation in agricultural research. Rather, generating useful results *depends* on the large-scale participation of diverse, independent contributors. Because data are generated de-centrally by farmers, but analysis of aggregated data is carried out by researchers, a key question relates to the collection of data and feedback of results. Live phone calls by paid enumerators can be an option, but this likely becomes costly with high numbers of participants (see Dillon 2012). Beza et al. (2018) suggest that the use of SMS for transmitting farmer-generated data is feasible where contributors are clearly aware of personal benefits associated with data provision. Automated calls using interactive voice response (IVR), where data are entered through the phone's number pad, can be another alternative, especially for illiterate farmers (Chancellor et al. 2019). The use of IVR requires some training to ensure data quality (Patnaik et al. 2009; Lerer et al. 2010). But in practical applications of tricot research, farmers receive at least one initial training about planting and observing the tricot trials, anyway (Steinke and Van Etten 2016). Training units on systematic reporting through IVR or SMS could also be included there.

In the future, agricultural citizen science, using the tricot experimental format, need not be limited to the evaluation of crop varieties. Triadic comparisons of different technology variants may be used to generate evidence on a large variety of agricultural technologies, for example, different tillage styles, soil amendments, or poultry breeds. As was demonstrated, the 'Wisdom of Crowds' principle implies that existing qualitative differences between the technology variants can be detected with a sufficiently high number of observations, as long as these differences have practical significance. In some cases, where such differences are small or challenging to spot, potentially large needs for farmer training may play in favor of choosing stronger scientist supervision, while 'simple' observations can be crowdsourced more easily (see Oliver et al. 2019).

In summary, the presented evidence on the accuracy of farmer-generated data in tricot-style research suggests that tricot can be a viable tool for involving large numbers of farmers in knowledge generation through agricultural on-farm experiments. Because the usefulness of results generally increases with the number of contributors, researchers are incentivized to cooperate with extension services or community organizations to achieve truly large-scale farmer participation. It is clear that modern ICT play a key role for the communication of raw data (on-farm observations) and processed results between many farmers and few researchers. But how this should be done in practice – e.g., through live phone calls, SMS, or automated calls – is not determined yet and will likely depend on context, including the degree of farmers' trust to the research organization, their perception of benefits, and their technology use habits (Steinke and Van Etten 2016; Wyche and Steinfield 2016; Beza et al. 2018).

Research question 2

How can the Positive Deviance approach help to identify locally suitable innovation using ICT-mediated data inputs from many smallholder farmers?

Positive Deviance emerged in the 1990s as an action research approach to identify promising childcare and nutrition behavior in Asia and Africa (Sternin et al. 1998; Marsh et al. 2004; Bisits Bullen 2011). Over the years, the methodology has been applied by research and development organizations to a diversity of intervention areas, including HIV prevention, hospital patient safety, and pregnancy in resource-poor environments (Ahrari et al. 2002; Lapping et al. 2002; Lawton et al. 2014). Implementations of the Positive Deviance approach typically rely on (i) the systematic analysis of quantitative data about research subjects, and (ii) qualitative follow-up inquiry about the behaviors leading to the desired outcomes.

In the field of agriculture, however, existing scientific studies of positive deviant behavior have generally taken a more explorative, anecdotal approach (Ochieng 2007; Biggs 2008). In addition, definitions of what constitutes Positive Deviance in agricultural development vary, given the multi-functionality of agriculture. Consequently, to determine how more systematic applications of the Positive Deviance approach can contribute to agricultural development, three points need to be addressed: First, a definition of positive deviant outcomes is needed that accounts for the multi-functionality of agriculture, avoiding strong trade-offs between competing goals of smallholder farming. Second, there is a need to define how quantitative data on smallholder agriculture can be used to identify positive deviants. And third, it must be determined whether these positive deviants can indeed be models for locally viable agricultural development and thus inform the local prioritization of interventions.

Recently, Modernel et al. (2018) illustrated the use of multiple household-level indicators for identifying *Pareto-optimal* household performance regarding these different dimensions. Pareto-optimality is a useful concept to approach overall household performance in smallholder agriculture, which must address a number of needs simultaneously (Tittonell et al. 2007a; Groot et al. 2012; Klapwijk et al. 2014). Defining Positive Deviance by the Pareto-optimal relative achievement of multiple agricultural development goals avoids favoring any particular farming style and accounts for possible trade-offs between different objectives. Nonetheless, viable implementations of the Positive Deviance approach in agriculture necessarily imply choosing a limited number of goals to be included. This may mean that important positive of negative outcomes of agriculture remain outside the picture. For example, the implementation presented here (Steinke et al. 2019) does not include information on long-term resource management and economic farm viability (see Tittonell et al. 2007a).

Through qualitative follow-up interviews with 15 identified positive deviants, the study presented here identified 14 locally viable practices as options for multi-objective agricultural development. This supports the idea that an adaptation of the Positive Deviance approach, which selects study households from a Pareto-front on multiple dimensions of agricultural development, can generate useful inputs to local innovation process. So far, it is unclear how agricultural development interventions emphasizing the identified options would affect adoption dynamics compared to more conventional development programs that introduce innovation previously validated *ex-situ*.

An important question for the practical feasibility of applying the Positive Deviance approach for agricultural development relates to the resource-intensity of household data enumeration. Large-scale survey efforts, as implemented for the proof-of-concept presented here, are sometimes carried out by public and private extension providers as part of their ongoing monitoring and evaluation activities. These data could be used for identifying positive deviants. Alternatively, household surveys could be recurrently administered by extension staff at specific occasions, such as training events or farm visits. Under this scenario, the use of survey software on personal smartphones of advisory staff is likely to facilitate standardized data collection and continuous compilation (Hartung et al. 2010).

In the future, mobile phone-mediated agro-advisory applications may help to avoid the need for costly enumeration. Two-way communication between farmers and digital services offers new opportunities for autonomous data entry by farmers. 'Information pull' services, where registered farmers use their phones to access relevant information, could ask farmers a small set of different questions each time they access the service, using technologies such as IVR, SMS, or USSD codes. Over time, as the partial questionnaires generate data, individual user profiles can fill up. This way, a comprehensive household survey could be completed across an extended period of time, such as an agricultural season. If the digital advisory application uses the farmer-provided data for targeting different contents to different users (see Section 5.3), farmers have an incentive to provide accurate data. In practice, additional incentives may be needed to encourage sufficiently frequent usage of this kind of service. But the provision of mobile pay-outs or other benefits to farmer-users is still likely to imply lower costs than face-to-face survey enumeration by agricultural extension staff.

Research question 3

How can digital two-way communication be employed to target agro-advisory messages to heterogeneous smallholder farmers?

Smallholder context is highly diverse regarding farming systems, market opportunities, climatic conditions, and other aspects (e.g., Tittonell et al. 2010; Pacini et al. 2014; Kuivanen et al. 2016b). Along with rapidly changing pressures and societal demands to smallholder agriculture, as well as heterogeneous preferences and aspirations, this diversity implies individual farmers can have strongly differentiated information needs. As a consequence, digital agro-advisory applications employed by extension services require a large body of advisory messages to adequately address specific information needs. Through their personal mobile phones, farmers may express specific interests and request preferred information. But many additional factors, beyond individual preferences, may warrant disaggregation of advisory contents, such as the farmer's geographical location, labor availability, or disposable capital (Muthoni et al. 2017; Descheemaeker et al. 2019). Therefore, automated two-way communication can help to match supply and demand of agricultural advice. Digital agro-advisory applications could let farmers provide data about their farm or household, and then use this data to select advisory messages accordingly.

Data collection through mobile phones can take many forms, however. In practice, requesting users to answer extensive questionnaires through an automated interface, such as IVR, to obtain a comprehensive household characterization is not likely to be feasible due to user fatigue (Kilic and Sohnesen 2019). To decide how two-way communication can contribute to viable targeting of agro-advisory messages in practice, a mechanism is needed that (i) requires only limited data inputs from farmers, (ii) is convenient and simple to use, ensuring data accuracy, and (iii) generates information that can be used for useful individual prioritization of messages.

One way to achieve such a viable minimum-data approach to targeting agro-advisory consists in enumerating selected indicators from farmers that are highly predictive of their individual information needs. The work presented here (Section 5.3) uses standard indicators of household characterization as predictors of farmers' information preferences (as an approximation of information *needs*). The results suggest that small sets of two to four indicators can be sufficient to prioritize a set of agro-advisory messages in a way that reduces the risk of delivering predominantly irrelevant information. Farmers would need to answer between five and ten simple questions to provide that information. Using the established RHOMIS question format implied that questions were easy to understand and generated accurate responses (Hammond et al. 2017; Fraval et al. 2019).

The suggested solution is, however, but one possible implementation of the trade-off between a low data entry effort and high household specificity of advice. Practical tests with actual digital two-way communication interfaces now need to explore whether farmers will perceive the data entry effort as justified by the improved customization of advice. If needed, the extent of data enumeration may also be reduced, albeit likely at the cost of reduced quality of targeting. A technically more advanced solution could consist in splitting up the enumeration of household data, asking farmers a different sub-set of a larger pool of questions each time they access the agro-advisory application.

A crucial input for generating household-specific prioritizations of advisory messages, however, consists in creating the link between enumerated household characteristics and corresponding information needs. In the small-scale proof-of-concept study presented here, farmers' information needs were approximated by their information preferences expressed in a choice experiment. For future agro-advisory applications, however, fitting predictive models using preference data collected by classical stated preference methods will not likely be a viable strategy, given the associated costs of enumeration, as well as possible mismatches between stated preferences and (evolving) actual information needs. An important challenge for future digital advisory services will consist in harnessing the potential of analyzing the preference choices made by users *within* the application itself. That is, farmers may begin with using an information service that provides only weakly targeted agricultural advice. But in addition to providing household data, users also generate data by selecting certain contents and disregarding others within the application.

Continuously overlaying this accumulating data on users' choices and user characteristics allows creating and recurrently adapting algorithms that create increasingly customized suggestions for individual users. Such big data algorithms could create powerful, adaptive recommendation systems. Resulting suggestions might take the form of "Other farmers from your region who were interested in topic X have also listened to message Y." These types of recommendation systems are already implemented by digital content platforms such as Youtube (for videos), Spotify (for music), or Google Feed (for third-party news articles). Here, despite tremendous numbers of options, algorithms make customized suggestions based on the user's personal access history and known characteristics, such as gender, age, location, or language.

Direct network effects imply that algorithm-based recommendation systems improve with the number and diversity of active users (Gawer 2014). Therefore, rather than creating isolated digital services for specific topics, extension service may more successfully invest into the creation of integrated information services that can satisfy a diversity of knowledge and information needs. In the future, agricultural advisory applications might even be integrated with efforts of 'mHealth' or 'eHealth' (using ICT to deliver medical services in remote areas, see Kahn et al. 2010; Chib et al. 2015). Joining these efforts could be a strategy to capitalize on public investments into digital development, as well as increasing and diversifying the user base by offering 'all-in-one' information services for rural citizens in developing countries.

6.2 Implications for extension services

Employing large-scale farmer participation through the use of modern ICT has the potential to improve the performance of extension services in various ways. All approaches explored in this dissertation involve the collection of data from many individual farmers via digital tools (in particular, mobile phones), processing of aggregated data by extension services, and subsequent feedback of selected information to individual clients. This approach to agricultural extension brings about a number of challenges under current conditions. Most notably, the reliance on farmer-generated data inputs stands in contrast to a more conventional focus on 'expert' knowledge, which may imply fundamental shifts in decision-making processes within extension services: In agricultural citizen science, data generated by farmers, rather than by researcher-led trials, are used to assess technology suitability. The Positive Deviance approach implies using farmers' data inputs in order to identify households with outstanding performance and superior practice, rather than extension services selecting or appointing 'model farms' and implementing a pre-defined technology transfer agenda. And digital two-way communication can give farmers access to a high diversity of different information and technology packages, rather than emphasizing a low number of established priority technologies.

Consequently, harnessing the potential of large-scale farmer participation through modern ICT is, in many cases, likely to require changes in organizational culture and institutions within advisory providers. Despite the progress towards decentralization that was made in many countries over the last decades, top-down agenda-setting by (local) extension managers is still the rule (Davis 2008). A greater use of the concepts explored in this dissertation would potentially reallocate decision-making power: from the management level of extension services towards field agents (e.g., identifying positive deviant practices on-site) and technological systems (e.g., algorithmic recommendation systems). In practice, such redistribution of inter-organizational power needs to be accompanied by adequate training at all levels: Senior management must be enabled to monitor and report on diverse, data-driven, and sometimes unforeseen advisory activities, such as the promotion of farmer-bred crop varieties based on citizen science data (van Etten et al. 2019a). Field agents, on the other hand, might require new abilities for data analysis and increased autonomous decision-making based on farmer-provided and crowdsourced data (Heeks 2002; Janssen et al. 2017; Müller et al. 2018).

Recent experiences demonstrate that effective large-scale farmer participation in agricultural extension is likely to depend on the availability of user-friendly digital applications that facilitate the collection, integration, and/or analysis of farmers' contribution. Existing examples include websites for experimental design, IVR systems for data collection, and smartphone applications for advisory delivery using two-way communication (Carmona et al. 2018; van Etten et al. 2018; Eitzinger et al. 2019). User-centered, iterative design is needed to ensure these applications comply with the communication needs, habits, and preferences of users (farmers and extension staff alike). But even so, new digital agro-advisory applications risk falling out of use if there is limited capacity within extension organizations for adapting, updating, developing, and troubleshooting the service (Heeks 2002). Consequently, sustaining digital services for agricultural advisory will require investments into staff training and capacity-building for digital design and related fields, such as statistics and data science.

6.3 Limitations and further research needs

This dissertation addressed questions relevant to the feasibility of engaging large-scale farmer participation in agricultural extension through digital tools. Evidence was generated through three independent proof-of-concept studies. This helped to judge the potential of addressing specific, existing limitations of agricultural extension by employing agricultural citizen science, the Positive Deviance approach, and digital two-way communication. The derived conclusions on feasibility may have important limitations, however. Each empirical test presented in Section 5 was performed at one point in time and at few, selected locations, only. As a consequence, conclusions on feasibility might be challenged under different circumstances. For example, the number of farmers needed to reliably distinguish a set of crop varieties with the tricot crowdsourcing methodology (Section 5.1) may vary by region, crop, varietal portfolio, seasonal climate, and other factors. Although the usefulness of the tested approaches has been demonstrated in principle, further research efforts are needed to adapt practical implementations to local context. This may imply local replications of the research procedures taken here, for example, to define locally meaningful disaggregating household variables for efficient two-way communication in agro-advisory applications.

Another possible limitation relates to the limited involvement of active advisory staff in the development and testing of the new approaches presented here. Since experimental design and data analysis were carried out by researchers, it is yet unclear how the new concepts can be integrated with existing work routines at extension services. Further research needs to explore the respective perceptions, needs, and capacities of extension staff, in order to determine how large-scale farmer participation through modern ICT can effectively enhance the performance of agricultural extension services (see Tata and McNamara 2016; Birke et al. 2019).

The evidence presented here allows making statements about the potential of engaging large numbers of farmers through modern ICT. Actual development of digital applications, however, would have been too time- and cost-intensive for the purpose of the research presented here. Consequently, farmers' interactions with ICT was largely emulated, and data was recorded by field enumerators. This constraint has important implications. While the explored concepts are promising under the assumption that farmers use ICT to provide data, the conditions under which this assumption holds true may vary greatly by context. Highly participatory research and user-centered design processes are now needed to create digital applications that smallholder farmers will engage with actively (Carter and Hundhausen 2010; Pitula and Radhakrishnan 2011; Wyche and Steinfield 2016).

Lastly, the research presented here emphasizes the conceptual feasibility of improving agricultural extension by engaging farmers' contributions. Despite the potential of the tested concepts, however, it is not clear to what extent their application may improve outcomes, compared to a counterfactual of business-as-usual agricultural advisory. For the case of agricultural citizen science, recent research suggests it enables increased geographic fine-tuning and diversification of crop varietal recommendations, resulting in reduced climate risk for smallholder farmers (van Etten et al. 2019a; Fadda and Van Etten 2019). Especially for the other two concepts, however, further research is needed to understand possible contributions to improved performance of agricultural extension. In the future, randomized control trials and panel survey analysis could help to evaluate the effects of applying the Positive Deviance approach or digital two-way communication on technology adoption, farm performance, and rural household welfare.

7 Outlook

Over the last two decades, private and public agricultural extension services in developing countries have created a diversity of digital information services targeted at smallholder farmers. Yet only few digital advisory applications have experienced significant uptake beyond the piloting phase (Qiang et al. 2012). In addition to other reasons, such as insufficient client orientation, limited uptake may also be related to a growing fragmentation of the digital service landscape: Many agro-advisory applications focus exclusively on highly specified domains of farming, for example, by providing technical advice, price information, or market facilitation services, and typically for certain value chains, only (Baumüller 2018). For farmers, this means that each information service implies learning costs, which may not be trivial and are associated with uncertain benefits (Wyche and Steinfield 2016). The limited scope of many applications also disincentivizes extension services to widely advertise an agro-advisory application, as it may be relevant to a limited share of farmers, only. A key challenge for the future development of ICT applications will, therefore, likely consist in the integration of a number of diverse information services within a single application.

Such a one-stop application could require considerable public or private investment. But offering different types of information services in one product – e.g., production advice, facilitation of market linkages, and farmer-to-farmer exchange of experiences – has the potential to attract a large and diverse range of users. Through direct network effects, the increasing number and diversity of users would allow a digital agro-content platform to improve its overall service (Gawer 2014). If different kinds of two-way communication are built into all offered services, the accumulating household data can be used to make internal cross-linkages and improve the advisory service overall. Say, for example, at some point in the farming season, the users of the production advisory function within the application increasingly search for information on treating maize rust. All users, including those who concentrate on a second function for market linkages, could then receive a warning about an imminent rust outbreak. And the algorithm managing a third

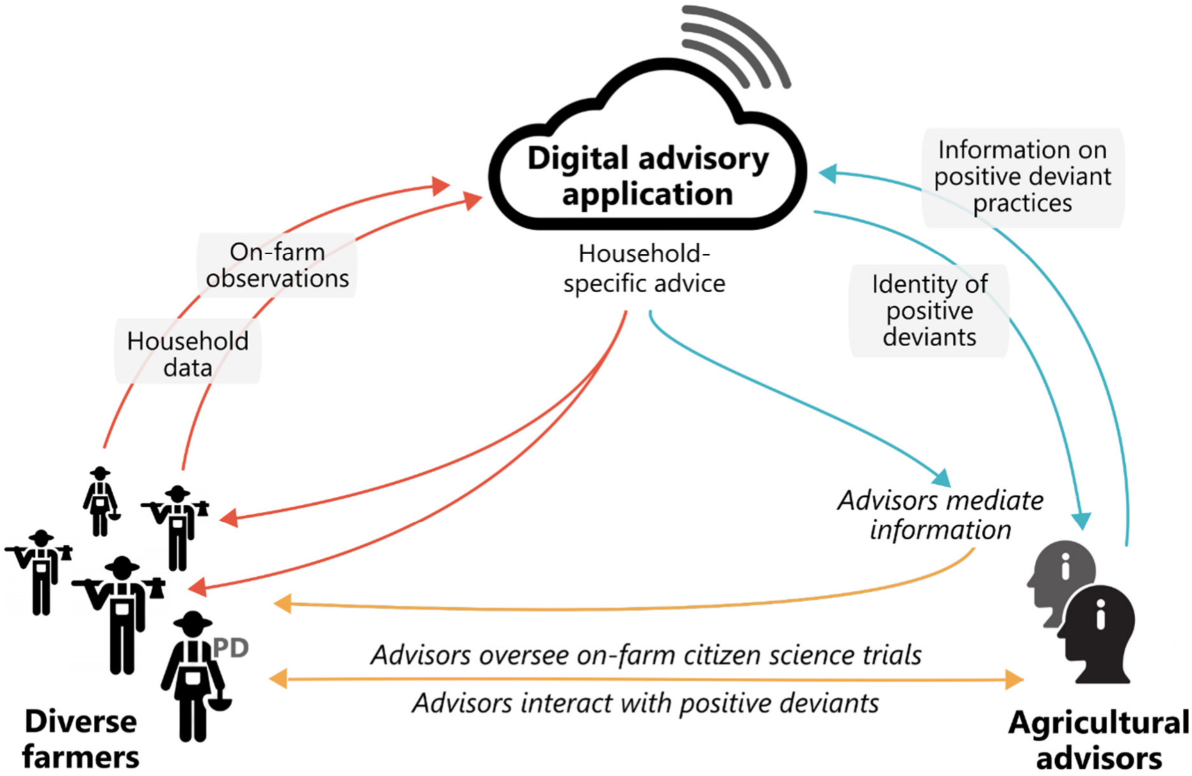

Figure 4: A model for integrating the three concepts explored in this dissertation within a single digital agro-advisory application. In red, information exchange between individual farmers and the service via two-way communication. In blue, information exchange between agricultural advisors and the service. Advisors may act as intermediaries, assisting farmers in using the service and helping with further interpretation and adaptation of household-specific advisory messages. In orange, activities required in face-to-face interaction between advisors and farmers.

function for farmer-to-farmer exchange could highlight farmer narrations about rust control and prevention. Many other ways of linking different information services through usage meta-data are conceivable. Allowing the users to use a versatile digital application in the partial way they prefer, while making suggestions and internal re-directions to other offered services within the application ecosystem, is part of the recipe for success applied by Facebook, the world market-dominating social network.

Despite increasing availability of mobile services and continued pressure to increase cost-efficiency, it is most likely that on-site visits, face-to-face interaction, and practical capacity-building through extension staff will remain cornerstones of agricultural advisory in the developing world (Sulaiman et al. 2012; Matous et al. 2015). Rather than replacing advisors, ICT applications have great potential to enhance the quality of real-life farmer-advisor interactions. A number of digital services explicitly designed for use by extension agents is already available (Gandhi et al. 2009; Saito et al. 2015; Tata and McNamara 2016; Wright et al. 2016). With increasing integration of data-driven services into singly, potentially national agro-advisory applications, the scope of duties of extension agents may also involve acting as 'digital facilitators', assisting farmers in using the application. A well-designed application, however, might also allow tech-savvy rural youth to perform as information intermediaries, possible creating new job opportunities (Fu and Akter 2016; van Campenhout 2017; Bentley et al. 2019). This could free resources for actual extension staff to take up important roles in interpreting and further contextualizing the recommendations made by the application, as well as in knowledge (co-)creation, for example, by supervising crowdsourced on-farm experimentation.

As an illustration of the prospects outlined here – increasing integration of services, digital facilitation by extension agents –, Figure 4 suggests how the three concepts explored in this dissertation could be integrated into a single digital service. Such a service would use two-way communication interfaces to collect different types of data inputs from farmers, such as observations from small on-farm experiments, household resource endowments, farming activities, and outcomes. The service could then use this information (i) to generate evidence on the local suitability of agricultural technologies, (ii) to identify 'positive deviants', and (iii) to tailor the selection of advisory messages to the individual user (cf. Sections 5.1 – 5.3). Farmers could receive, for example, context-adapted recommendations on farm design, including customized crop varietal portfolios, as well as individually selected suggestions on positive deviant practices. Agricultural advisors might assist farmers in using the application, supervise the implementation of on-farm trials, and interact periodically with identified 'positive deviants' to identify superior practices that could later be promoted by the digital application.

To move from isolated solutions towards integrated digital information services, policy-makers can provide vital support. For example, with public recognition of the significance of digital communication and access to information for smallholder farming, policy-makers might introduce tax breaks or subsidies for the development of new agro-advisory applications and services. This could create opportunities for new business models around agricultural advisory, attracting private investment and fueling research and innovation. Local 'tech hubs' or 'incubators' can provide a physical space where digital designers, developers, and donors meet, exchange ideas, and lay foundations for collaboration (Qiang et al. 2012; Kelly and Firestone 2015). By providing these spaces, governments can support the emergence of scalable, financially sustainable, and interoperable digital agro-advisory services.

8 References

Affholder, F., D. Jourdain, D. D. Quang, T. P. Tuong, M. Morize, and A. Ricome. 2010. Constraints to farmers' adoption of direct-seeding mulch-based cropping systems: A farm scale modeling approach applied to the mountainous slopes of Vietnam. *Agricultural Systems* 103: 51–62.

Ahrari, M., A. Kuttab, S. Khamis, A. A. Farahat, G. L. Darmstadt, D. R. Marsh, and F. J. Levinson. 2002. Factors associated with successful pregnancy outcomes in Upper Egypt: A positive deviance inquiry. *Food and Nutrition Bulletin* 23: 83–88.

Aker, J. C. 2011. Dial "A" for agriculture: A review of information and communication technologies for agricultural extension in developing countries. *Agricultural Economics* 42: 631–647.

Aker, J. C., and I. M. Mbiti. 2010. Mobile phones and economic development in Africa. *Journal of Economic Perspectives* 24: 207–232.

Aker, J. C., I. Ghosh, and J. Burrell. 2016. The promise (and pitfalls) of ICT for agriculture initiatives. *Agricultural Economics* 47: 35–48.

Almekinders, C. J. M., K. Beumer, M. Hauser, M. Misiko, M. Gatto, A. O. Nkurumwa, and O. Erenstein. 2019. Understanding the relations between farmers' seed demand and research methods: The challenge to do better. *Outlook on Agriculture* 48: 16-21.

Anderson, J. R., and G. Feder. 2004. Agricultural extension: Good intentions and hard realities. *World Bank Research Observer* 19: 41–60.

Andersson, J. A., and S. D'Souza. 2014. From adoption claims to understanding farmers and contexts: A literature review of Conservation Agriculture (CA) adoption among smallholder farmers in southern Africa. *Agriculture, Ecosystems and Environment* 187: 116–132.

Arslan, A., N. Mccarthy, L. Lipper, S. Asfaw, and A. Cattaneo. 2014. Adoption and intensity of adoption of conservation farming practices in Zambia. *Agriculture, Ecosystems and Environment* 187: 72–86.

Asfaw, S., B. Shiferaw, F. Simtowe, and L. Lipper. 2012. Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. *Food Policy* 37: 283–295.

Ashby, J. A., A. R. Braun, T. Gracia, M. del Pilar Guerrero, C. A. Quirós, and J. I. Roa. 2000. Investing in farmers as researchers. Experience with local agricultural research committees in Latin America. CIAT publication. Vol. 318. Cali, Colombia: CIAT.

Barrett, C. B. 2008. Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy* 33: 299–317.

Baudron, F., M. Misiko, B. Getnet, R. Nazare, J. Sariah, and P. Kaumbutho. 2019. A farm-level assessment of labor and mechanization in Eastern and Southern Africa. *Agronomy for Sustainable Development* 39: 17.

Baumüller, H. 2015. Assessing the role of mobile phons in offering price information and market linkages: The case of m-farm in Kenya. *The Electronic Journal of Information Systems in Developing Countries* 68: 1–16.

Baumüller, H. 2018. The little we know: An exploratory literature review on the utility of mobile phone-enabled services for smallholder farmers. *Journal of International Development* 154: 134–154.

Benson, A., and T. Jafry. 2013. The state of agricultural extension: An overview and new caveats for the future. *Journal of Agricultural Education and Extension* 19: 381–393.

Bentley, J. W., P. van Mele, N. F. Barres, F. Okry, and J. Wanvoeke. 2019. Smallholders download and share videos from the internet to learn about sustainable agriculture. *International Journal of Agricultural Sustainability* 17: 92–107.

Bernet, T., O. Ortiz, R. D. Estrada, R. Quiroz, and S. M. Swinton. 2001. Tailoring agricultural extension to different production contexts: A user-friendly farm-household model to improve decision-making for participatory research. *Agricultural Systems* 69: 183–198.

Berre, D., F. Baudron, M. Kassie, P. Craufurd, and S. Lopez-Ridaura. 2019. Different ways to cut a cake: Comparing expert-based and statistical typologies to target sustainable intensification technologies, a case-study in Southern Ethiopia. *Experimental Agriculture* 55(S1): 191–207.

Beza, E., J. Steinke, J. van Etten, P. Reidsma, C. Fadda, S. Mittra, P. Mathur, and L. Kooistra. 2017. What are the prospects for citizen science in agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers. *PLoS ONE* 12: e0175700.

Beza, E., P. Reidsma, P. M. Poortvliet, M. M. Belay, B. Sjors Bijen, and L. Kooistra. 2018. Exploring farmers' intentions to adopt mobile Short Message Service (SMS) for citizen science in agriculture. *Computers and Electronics in Agriculture* 151: 295–310.

Bezner Kerr, R., S. Snapp, M. Chirwa, L. Shumba, and R. Msachi. 2007. Participatory research on legume diversification with Malawian smallholder farmers for improved human nutrition and soil fertility. *Experimental Agriculture* 43: 437–453.

Biggs, S. 2008. Learning from the positive to reduce rural poverty and increase social justice: Institutional innovations in agricultural and natural resources research and development. *Experimental Agriculture* 44: 37–60.

Birke, F. M., M. Lemma, and A. Knierim. 2019. Perceptions towards information communication technologies and their use in agricultural extension: case study from South Wollo, Ethiopia. *Journal of Agricultural Education and Extension* 25: 47–62.

Birner, R., K. Davis, J. Pender, E. Nkonya, P. Anandajayasekeram, J. Ekboir, A. Mbabu, D. J. Spielman, D. Horna, S. Benin, and M. Cohen. 2009. From best practice to best fit: A framework for designing and analyzing pluralistic agricultural advisory services worldwide. *Journal of Agricultural Education and Extension* 15: 341–355.

Bisits Bullen, P. A. 2011. The positive deviance/hearth approach to reducing child malnutrition: Systematic review. *Tropical Medicine and International Health* 16: 1354–1366.

Brodbeck, F. C., R. Kerschreiter, A. Mojzisch, D. Frey, and S. Schulz-Hardt. 2002. The dissemination of critical, unshared information in decision-making groups: The effects of pre-discussion dissent. *European Journal of Social Psychology* 56: 35–56.

Bryan, E., C. Ringler, B. Okoba, C. Roncoli, S. Silvestri, and M. Herrero. 2013. Adapting agriculture to climate change in Kenya: Household and community strategies and determinants. *Journal of Environmental Management* 114: 26-35.

van Campenhout, B. 2017. There is an app for that? The impact of community knowledge workers in Uganda. *Information, Communication & Society* 20: 530-550.

Carmona, M. A., F. J. Sautua, O. Pérez-Hernández, and J. I. Mandolesi. 2018. AgroDecisor EFC: First Android™ app decision support tool for timing fungicide applications for management of late-season soybean diseases. *Computers and Electronics in Agriculture* 144: 310–313.

Carter, A. S., and C. D. Hundhausen. 2010. How is user interface prototyping really done in practice? A survey of user interface designers. *In:* 2010 IEEE Symposium on Visual Languages and Human-Centric Computing, 207–211.

Ceccarelli, S., E. P. Guimarães, and E. Weltzien (eds.). 2009. Plant breeding and farmer participation. Rome: Food and Agriculture Organization of the United Nations.

Challinor, A. J., J. Watson, D. B. Lobell, S. M. Howden, D. R. Smith, and N. Chhetri. 2014. A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change* 4: 287–291.

Chancellor, T. C. B., J. S. H. Priebe, and P. A. Mkenda. 2019. Crowdsourcing field observations from smallholder farmers in Tanzania using Interactive Voice Response. *Outlooks on Pest Management* 30: 104-110.

Chib, A., M. H. van Velthoven, and J. Car. 2015. mHealth adoption in low-resource environments: A review of the use of mobile healthcare in developing countries. *Journal of Health Communication* 20: 4–34.

Classen, L., S. Humphries, J. Fitzsimons, S. Kaaria, J. Jiménez, F. Sierra, and O. Gallardo. 2008. Opening participatory spaces for the most marginal: Learning from collective action in the Honduran hillsides. *World Development* 36: 2402–2420.

Cock, J., T. Oberthür, C. Isaacs, P. Roman, A. Palma, J. Carbonell, J. Victoria, G. Watts, A. Amaya, L. Collet, G. Lema, and E. Anderson. 2011. Crop management based on field observations: Case studies in sugarcane and coffee. *Agricultural Systems* 104: 755–769.

Cohn, A. S., P. Newton, J. D. B. Gil, L. Kuhl, L. Samberg, V. Ricciardi, J. R. Manly, and S. Northrop. 2017. Smallholder agriculture and climate change. *Annual Review of Environment and Resources* 42: 347–375.

Conrad, C. C., and K. G. Hilchey. 2011. A review of citizen science and community-based environmental monitoring: Issues and opportunities. *Environmental Monitoring and Assessment* 176: 273–291.

Cooper, S., F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, and Foldit players. 2010. Predicting protein structures with a multiplayer online game. *Nature* 466: 756–760.

Corbeels, M., J. de Graaff, T. H. Ndah, E. Penot, F. Baudron, K. Naudin, N. Andrieu, G. Chirat, J. Schuler, I. Nyagumbo, L. Rusinamhodzi, K. Traore, H. D. Mzoba, and I. S. Adolwa. 2013. Understanding the impact and adoption of conservation agriculture in Africa: A multi-scale analysis. *Agriculture, Ecosystems and Environment* 187: 155–170.

Cornwall, A. 2003. Whose voices? Whose choices? Reflections on gender and participatory development. *World Development* 31: 1325–1342.

Crane, T. A. 2010. Of models and meanings : Cultural resilience in social–ecological systems. *Ecology and Society* 15: 19.

Darnhofer, I., S. Bellon, B. Dedieu, and R. Milestad. 2010. Adaptiveness to enhance the sustainabiltiy of farming systems. *Agronomy for Sustainable Development* 30: 545–555.

Daum, T., H. Buchwald, A. Gerlicher, and R. Birner. 2018. Smartphone apps as a new method to collect data on smallholder farming systems in the digital age: A case study from Zambia. *Computers and Electronics in Agriculture* 153: 144–150.

Davis, K. 2008. Extension in Sub-Saharan Africa: Overview and assessment of past and current models, and future prospects. *Journal of International Agricultural and Extension Education* 15: 15–28.

Davis, K., E. Nkonya, E. Kato, D. A. Mekonnen, M. Odendo, R. Miiro, and J. Nkuba. 2012. Impact of Farmer Field Schools on agricultural productivity and poverty in East Africa. *World Development* 40: 402–413.

Deichmann, U., A. Goyal, and D. Mishra. 2016. Will digital technologies transform agriculture in developing countries? *Agricultural Economics* 47: 21–33.

Descheemaeker, K., E. Ronner, M. Ollenburger, A. C. Franke, C. J. Klapwijk, G. N. Falconnier, J. Wichern, and K. E. Giller. 2019. Which options fit best? Operationalizing the socio-ecological niche concept. *Experimental Agriculture* 55(S1): 169–190.

Dickinson, J. L., J. Shirk, D. Bonter, R. Bonney, R. L. Crain, J. Martin, T. Phillips, and K. Purcell. 2012. The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment* 10: 291–297.

Dillon, B. 2012. Using mobile phones to collect panel data in developing countries. *Journal of International Development* 24: 518–527.

Dorward, A., S. Anderson, Y. N. Bernal, E. S. Vera, J. Rushton, J. Pattison, and R. Paz. 2009. Hanging in, stepping up and stepping out: Livelihood aspirations and strategies of the poor. *Development in Practice* 19: 240–247.

Duncombe, R. 2015. Mobile phones for agricultural and rural development: A literature review and suggestions for future research. *European Journal of Development Research* 28: 213-235.

Eiben, C. B., J. B. Siegel, J. B. Bale, S. Cooper, F. Khatib, B. W. Shen, Foldit Players, B. L. Stoddard, Z. Popović, and D. Baker. 2012. Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nature Biotechnology* 30: 190–192.

Eitzinger, A., J. Cock, K. Atzmanstorfer, C. R. Binder, P. Läderach, O. Bonilla-Findji, M. Bartling, C. Mwongera, L. Zurita, and A. Jarvis. 2019. GeoFarmer: A monitoring and feedback system for agricultural development projects. *Computers and Electronics in Agriculture* 158: 109–121.

van Etten, J. 2011. Crowdsourcing crop improvement in Sub-Saharan Africa: A proposal for a scalable and inclusive approach to food security. *IDS Bulletin* 42: 102–110.

van Etten, J., C. Quirós, B. Madriz, and K. de Sousa. 2018. Climmob, digital platform to support triadic comparisons of technologies (tricot). CGIAR Gender & Breeding Initiative.

van Etten, J., K. de Sousa, A. Aguilar, M. Barrios, A. Coto, M. Dell'Acqua, C. Fadda, Y. Gebrehawaryat, J. van de Gevel, A. Gupta, A. Y. Kiros, B. Madriz, P. Mathur, D. K. Mengistu, L. Mercado, J. N. Mohammed, A. Paliwal, M. E. Pè, C. F. Quirós, J. C. Rosas, N. Sharma, S. S. Singh, I. S. Solanki, and J. Steinke. 2019a. Crop variety management for climate adaptation supported by citizen science. *PNAS* 116: 4194–4199.

van Etten, J., E. Beza, L. Calderer, K. van Duijvendijk, C. Fadda, B. Fantahun, Y. G. Kidane, J. van de Gevel, A. Gupta, D. K. Mengistu, D. Kiambi, P. N. Mathur, L. Mercado, S. Mittra, M. J. Mollel, J. C. Rosas, J. G. Suchini, and K. S. Zimmerer. 2019b. First experiences with a novel farmer citizen science approach: Crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture* 55(S1): 275-296.

Fadda, C., and J. van Etten. 2019. Generating farm-validated variety recommendations for climate adaptation. *In:* T. Rosenstock, E. Girvetz, and A. Nowak (eds.). The Climate-Smart Agriculture Papers, 127–138. Cham: Springer.

Farrington, J., and A. M. Martin. 1988. Farmer participatory research: A review of concepts and recent fieldwork. *Agricultural Administration and Extension* 29: 247–264.

Faure, G., Y. Desjeux, and P. Gasselin. 2012. New challenges in agricultural advisory services from a research perspective: A literature review, synthesis and research agenda. *Journal of Agricultural Education and Extension* 18: 461–492.

Feder, G., R. E. Just, and D. Zilberman. 1985. Adoption of agricultural innovations in developing countries: A Survey. *Economic Development and Cultural Change* 33: 255–298.

Feder, G., R. H. Slade, and A. K. Sundaram. 1986. The Training and Visit extension system: An analysis of operations and effects. *Agricultural Administration* 21: 33–59.

Feder, G., J. R. Anderson, R. Birner, and K. Deininger. 2010. Promises and realities of community-based agricultural extension. IFPRI Discussion Paper 00959. Washington D.C., USA: IFPRI.

van de Fliert, E., J. Pontius, and N. Röling. 1995. Searching for strategies to replicate a successful extension approach: Training of IPM trainers in Indonesia. *European Journal of Agricultural Education and Extension* 1: 41–63.

Fraval, S., J. Hammond, J. Wichern, S. J. Oosting, I. J. M. De Boer, N. Teufel, M. Lannerstad, K. Waha, T. Pagella, T. S. Rosenstock, K. E. Giller, M. Herrero, D. Harris, and M. T. van Wijk. 2019. Making the most of imperfect data: A critical evaluation of standard information collected in farm household surveys. *Experimental Agriculture* 55: 230-250.

Frelat, R., S. Lopez-Ridaura, K. E. Giller, M. Herrero, S. Douxchamps, A. Andersson Djurfeldt, O. Erenstein, B. Henderson, M. Kassie, B. K. Paul, C. Rigolot, R. S. Ritzema, D. Rodriguez, P. J. A. van Asten, and M. T. van Wijk. 2016. Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *PNAS* 113: 458–463.

Friis-Hansen, E., and D. Duveskog. 2012. The empowerment route to well-being: An analysis of Farmer Field Schools in East Africa. *World Development* 40: 414–427.

Fu, X., and S. Akter. 2016. The impact of mobile phone technology on agricultural extension services delivery: Evidence from India. *Journal of Development Studies* 52: 1561–1576. doi:10.1080/00220388.2016.1146700.

Gandhi, R., R. Veeraraghavan, K. Toyama, and V. Ramprasad. 2009. Digital Green: Participatory video and mediated instruction for agricultural extension. *Information Technologies and International Development* 5: 1–15.

Gawer, A. 2014. Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy* 43: 1239–1249.

Gigerenzer, G., and W. Gaissmaier. 2011. Heuristic decision making. *Annual Review of Psychology* 62: 451–482.

Giller, K. E., E. Witter, M. Corbeels, and P. Tittonell. 2009. Conservation agriculture and smallholder farming in Africa: The heretics' view. *Field Crops Research* 114: 23–34.

Giller, K. E., M. Corbeels, J. Nyamangara, B. Triomphe, F. Affholder, E. Scopel, and P. Tittonell. 2011. A research agenda to explore the role of conservation agriculture in African smallholder farming systems. *Field Crops Research* 124: 468–472.

Glover, D., J. Sumberg, and J. A. Andersson. 2016. The adoption problem; or why we still understand so little about technological change in African agriculture. *Outlook on Agriculture* 45: 3-6.

Grabowski, P. P., J. M. Kerr, S. Haggblade, and S. Kabwe. 2016. Determinants of adoption and disadoption of minimum tillage by cotton farmers in eastern Zambia. *Agriculture, Ecosystems and Environment* 231: 54–67.

Groot, J. C. J., G. J. M. Oomen, and W. A. H. Rossing. 2012. Multi-objective optimization and design of farming systems. *Agricultural Systems* 110: 63–77.

GSM Association (GSMA). 2018. The Moblie Economy Sub-Saharan Africa 2018. London, UK: GSMA.

Hammond, J., S. Fraval, J. van Etten, J. G. Suchini, L. Mercado, T. Pagella, R. Frelat, M. Lannerstad, S. Douxchamps, N. Teufel, D. Valbuena, and Mark T. van Wijk. 2017. The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. *Agricultural Systems* 151: 225–233.

Hand, E. 2010. People power. *Nature* 466: 685–687.

Hansen, J., J. Hellin, T. Rosenstock, E. Fisher, J. Cairns, C. Stirling, C. Lamanna, J. van Etten, A. Rose, and B. Campbell. 2019. Climate risk management and rural poverty reduction. *Agricultural Systems* 172: 28–46.

Hartung, C., Y. Anokwa, W. Brunette, A. Lerer, C. Tseng, and G. Borriello. 2010. Open Data Kit: Tools to build information services for developing regions. Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development.

Hazell, P., and S. Wood. 2008. Drivers of change in global agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363: 495–515.

Heeks, R. 2002. Information systems and developing countries: Failure, success, and local improvisations. *The Information Society* 18: 101–112.

Henrich, J. 2001. Cultural transmission and the diffusion of innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change. *American Anthropologist* 103: 992–1013.

Hochachka, W. M., D. Fink, R. A. Hutchinson, D. Sheldon, W. Wong, and S. Kelling. 2012. Data-intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution* 27: 130–137.

Hoffmann, V., K. Probst, and A. Christinck. 2007. Farmers and researchers: How can collaborative advantages be created in participatory research and technology development? *Agriculture and Human Values* 24: 355–368.

Hoogeveen, J., K. Croke, A. Dabalen, G. Demombynes, and M. Giugale. 2014. Collecting high frequency panel data in Africa using mobile phone interviews. *Canadian Journal of Development Studies* 35: 186–207.

Howden, S. M., J.-F. Soussana, F. N. Tubiello, N. Chhetri, M. Dunlop, and H. Meinke. 2007. Adapting agriculture to climate change. *PNAS* 104: 19691–19696.

Humphreys, M., W. A. Masters, and M. E. Sandbu. 2006. The role of leaders in democratic deliberations: Results from a field experiment in São Tome and Principe. *World Politics* 58: 583–622.

Humphries, S., J. C. Rosas, M. Gómez, J. Jiménez, F. Sierra, O. Gallardo, C. Avila, and M. Barahona. 2015. Synergies at the interface of farmer–scientist partnerships: Agricultural innovation through participatory research and plant breeding in Honduras. *Agriculture & Food Security* 4: 27.

Hunter, J., A. Alabri, and C. van Ingen. 2013. Assessing the quality and trustworthiness of citizen science data. *Concurrency and Computation: Practice and Experience* 25: 454–466.

IPCC. 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories.

ITU. 2018. World Telecommunication/ICT Indicators database. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/publications/wtid.aspx

Janssen, S. J. C., C. H. Porter, A. D. Moore, I. N. Athanasiadis, I. Foster, J. W. Jones, and J. M. Antle. 2017. Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology. *Agricultural Systems* 155: 200–212.

Johnson, N. L., N. Lilja, and J. A. Ashby. 2003. Measuring the impact of user participation in agricultural and natural resource management research. *Agricultural Systems* 78: 287–306.

Jones, M., and F. Kondylis. 2018. Does feedback matter? Evidence from agricultural Services. *Journal of Development Economics* 131: 28–41.

Jones, P. G., and P. K. Thornton. 2003. The potential impacts of climate change on maize production in Africa and Latin America in 2055. *Global Environmental Change* 13: 51–59.

Joshi, K. D., K. P. Devkota, D. Harris, N. P. Khanal, B. Paudyal, A. Sapkota, and J. R. Witcombe. 2012. Participatory research approaches rapidly improve household food security in Nepal and identify policy changes required for institutionalisation. *Field Crops Research* 131: 40–48.

Kahn, J. G., J. S. Yang, and J. S. Kahn. 2010. "Mobile" health needs and opportunities in developing countries. *Health Affairs* 29: 254–260.

Kelly, T., and R. Firestone. 2015. Digital Dividends. How Tech Hubs are helping to Drive Economic Growth in Africa. Background paper to the World Development Report 2016 Digital Dividends. Washington, D.C., USA: World Bank.

Khatib, F., F. DiMaio, Foldit Contenders Group, Foldit Void Crushers Group, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska, I. Pichova, J. Thompson, Z. Popović, M. Jaskolski, and D. Baker. 2011. Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology* 18: 1175–1178.

Kilelu, C. W., L. Klerkx, C. Leeuwis, and A. Hall. 2011. Beyond knowledge brokering: An exploratory study on innovation intermediaries in an evolving smallholder agricultural system in Kenya. *Knowledge Management for Development Journal* 7: 84–108.

Kilelu, C. W., L. Klerkx, and C. Leeuwis. 2013. Unravelling the role of innovation platforms in supporting co-evolution of innovation: Contributions and tensions in a smallholder dairy development programme. *Agricultural Systems* 118: 65–77.

Kilic, T., and T. P. Sohnesen. 2019. Same question but different answer: Experimental evidence on questionnaire design's impact on poverty measured by proxies. *Review of Income and Wealth* 65: 144–165.

Klapwijk, C. J., M. T. van Wijk, T. S. Rosenstock, P. J. A. van Asten, P. K. Thornton, and K. E. Giller. 2014. Trade-offs in agricultural systems: Current status and way forward. *Current Opinion in Environmental Sustainability* 6: 110–115.

Knickel, K., G. Brunori, S. Rand, and J. Proost. 2009. Towards a better conceptual framework for innovation processes in agriculture and rural development: From linear models to systemic approaches. *Journal of Agricultural Education and Extension* 15: 131–146.

Knook, J., V. Eory, M. Brander, and D. Moran. 2018. Evaluation of farmer participatory extension programmes. *Journal of Agricultural Education and Extension* 24: 209–325.

Kremen, C., K. S. Ullman, and R. W. Thorp. 2011. Evaluating the quality of citizen-scientist data on pollinator communities. *Conservation Biology* 25: 607–617.

Kristjanson, P., R. S. Reid, N. Dickson, W. C. Clark, D. Romney, R. Puskur, S. MacMillan, and D. Grace. 2009. Linking international agricultural research knowledge with action for sustainable development. *PNAS* 106: 5047–5052.

Kuivanen, K. S., M. Michalscheck, K. Descheemaeker, S. Adjei-Nsiah, S. Mellon-Bedi, J. C. J. Groot, and S. Alvarez. 2016a. A comparison of statistical and participatory clustering of smallholder farming systems - A case study in Northern Ghana. *Journal of Rural Studies* 45: 184–198.

Kuivanen, K. S., S. Alvarez, M. Michalscheck, S. Adjei-Nsiah, K. Descheemaeker, S. Mellon-Bedi, and J. C. J. Groot. 2016b. Characterising the diversity of smallholder farming systems and their constraints and opportunities for innovation: A case study from the Northern Region, Ghana. *NJAS - Wageningen Journal of Life Sciences* 78: 153–166.

Lamers, D., M. Schut, L. Klerkx, and P. van Asten. 2017. Compositional dynamics of multilevel innovation platforms in agricultural research for development. *Science and Public Policy* 44: 739–752.

Lapping, K., D. R. Marsh, J. Rosenbaum, E. Swedberg, J. Sternin, M. Sternin, and D. G. Schroeder. 2002. The positive deviance approach: Challenges and opportunities for the future. *Food and Nutrition Bulletin* 23: 128–135.

Lawton, R., N. Taylor, R. Clay-Williams, and J. Braithwaite. 2014. Positive deviance: A different approach to achieving patient safety. *BMJ Quality and Safety* 23: 880–883.

Leeuwis, C. 2013. Communication for Rural Innovation. Rethinking Agricultural Extension. Oxford, UK: Blackwell Publishing.

Lerer, A., M. Ward, and S. Amarasinghe. 2010. Evaluation of IVR data collection UIs for untrained rural users. Proceedings of the First ACM Symposium on Computing for Development - ACM DEV ’10.

Lobell, D. B., W. Schlenker, and J. Costa-Roberts. 2011. Climate trends and global crop production since 1980. *Science* 333: 616–620.

Lopez-Ridaura, S., R. Frelat, M. T. Van Wijk, D. Valbuena, T. J. Krupnik, and M. L. Jat. 2018. Climate smart agriculture, farm household typologies and food security. An ex-ante assessment from Eastern India. *Agricultural Systems* 159: 57–68.

Lowder, S. K., J. Skoet, and T. Raney. 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development* 87: 16–29.

Lubell, M., M. Niles, and M. Hoffman. 2014. Extension 3.0: Managing agricultural knowledge systems in the network age. *Society and Natural Resources* 27: 1089–1103.

Mackintosh, U. A. T., D. R. Marsh, and D. G. Schroeder. 2002. Sustained positive deviant child care practices and their effects on child growth in Viet Nam. *Food and Nutrition Bulletin* 23: 16–25.

Marsh, D. R., D. G. Schroeder, K. A. Dearden, J. Sternin, and M. Sternin. 2004. The power of positive deviance. *BMJ* 329: 1177–1179.

Matous, P., Y. Todo, and A. Pratiwi. 2015. The role of motorized transport and mobile phones in the diffusion of agricultural information in Tanggamus Regency, Indonesia. *Transportation* 42: 771–790.

Mausch, K., D. Harris, E. Heather, E. Jones, J. Yim, and M. Hauser. 2018. Households' aspirations for rural development through agriculture. *Outlook on Agriculture* 47: 108–115.

McDonald, B. A., and E. H. Stukenbrock. 2016. Rapid emergence of pathogens in agro-ecosystems: Global threats to agricultural sustainability and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371: 20160026.

Meinzen-Dick, R., and M. Zwarteveen. 1998. Gendered participation in water management: Issues and illustrations from water users' associations in South Asia. *Agriculture and Human Values* 15: 337–345.

Méndez, V. E., M. Caswell, S. R. Gliessman, and R. Cohen. 2017. Integrating agroecology and participatory action research (PAR): Lessons from Central America. *Sustainability* 9: 1–19.

Minet, J., Y. Curnel, A. Gobin, J. P. Goffart, F. Mélard, B. Tychon, J. Wellens, and P. Defourny. 2017. Crowdsourcing for agricultural applications: A review of uses and opportunities for a farmsourcing approach. *Computers and Electronics in Agriculture* 142: 126–138.

Misiko, M. 2013. Dilemma in participatory selection of varieties. *Agricultural Systems* 119: 35–42.

Modernel, P., S. Dogliotti, S. Alvarez, M. Corbeels, V. Picasso, P. Tittonell, and W. A. H. Rossing. 2018. Identification of beef production farms in the Pampas and Campos area that stand out in economic and environmental performance. *Ecological Indicators* 89: 755–770.

Morton, J. F. 2007. The impact of climate change on smallholder and subsistence agriculture. *PNAS* 104: 19680–19685.

Moser, C. M., and C. B. Barrett. 2003. The disappointing adoption dynamics of a yield-increasing, low external-input technology: The case of SRI in Madagascar. *Agricultural Systems* 76: 1085–1100.

Müller, A., V. Mora, E. Rojas, J. Díaz, O. Fuentes, E. Giron, A. Gaytan, and J. van Etten. 2018. Emergency drills for agricultural drought response: a case study in Guatemala. *Disasters* 43: 410-430.

Müller, C., W. Cramer, W. L. Hare, and H. Lotze-Campen. 2011. Climate change risks for African agriculture. *PNAS* 108: 4313–4315.

Munthali, N., C. Leeuwis, A. Van Paassen, R. Lie, R. Asare, R. Van Lammeren, and M. Schut. 2018. Innovation intermediation in a digital age: Comparing public and private new-ICT platforms for agricultural extension in Ghana. *NJAS - Wageningen Journal of Life Sciences* 86: 64–76.

Muthoni, F. K., Z. Guo, M. Bekunda, H. Sseguya, F. Kizito, F. Baijukya, and I. Hoeschle-Zeledon. 2017. Sustainable recommendation domains for scaling agricultural technologies in Tanzania. *Land Use Policy* 66: 34–48.

Mwongera, C., K. M. Shikuku, J. Twyman, P. Läderach, E. Ampaire, P. van Asten, S. Twomlow, and L. A. Winowiecki. 2017. Climate smart agriculture rapid appraisal (CSA-RA): A tool for prioritizing context-specific climate smart agriculture technologies. *Agricultural Systems* 151: 192–203.

Nakasone, E., M. Torero, and B. Minten. 2014. The power of information: The ICT revolution in agricultural development. *Annual Review of Resource Economics* 6: 533–550.

Neef, A., and D. Neubert. 2011. Stakeholder participation in agricultural research projects: A conceptual framework for reflection and decision-making. *Agriculture and Human Values* 28: 179–194.

Nelson, R., R. Coe, and B. I. G. Haussmann. 2019. Farmer research networks as a strategy for matching diverse options and contexts in smallholder agriculture. *Experimental Agriculture* 55(S1): 125–144.

Newman, G., A. Wiggins, A. Crall, E. Graham, S. Newman, and K. Crowston. 2012. The future of citizen science: Emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment* 10: 298–304.

Niu, C., and C. Ragasa. 2018. Selective attention and information loss in the lab-to-farm knowledge chain: The case of Malawian agricultural extension programs. *Agricultural Systems* 165: 147–163.

Notenbaert, A., C. Pfeifer, S. Silvestri, and M. Herrero. 2017. Targeting, out-scaling and prioritising climate-smart interventions in agricultural systems: Lessons from applying a generic framework to the livestock sector in sub-Saharan Africa. *Agricultural Systems* 151: 153–162.

Ochieng, C. M. O. 2007. Development through Positive Deviance and its implications for economic policy making and public aministration in Africa: The case of Kenyan agricultural development, 1930-2005. *World Development* 35: 454–479.

Ojiem, J., N. de Ridder, B. Vanlauwe, and K. E. Giller. 2006. Socio-ecological niche: A conceptual framework for integration of legumes in smallholder farming systems. *International Journal of Agricultural Sustainability* 4: 79–93.

Olinto, P., K. Beegle, C. Sobrado, and H. Uematsu. 2013. The state of the poor: Where are the poor, where is extreme poverty harder to end, and what is the current profile of the world's poor? Economic Premise No. 125, October 2013. Washington, D.C., USA: The World Bank.

Oliver, K., A. Kothari, and N. Mays. 2019. The dark side of coproduction: Do the costs outweigh the benefits for health research? *Health Research Policy and Systems* 17: 33.

Ortiz, O., R. Orrego, W. Pradel, P. Gildemacher, R. Castillo, R. Otiniano, J. Gabriel, J. Vallejo, O. Torres, G. Woldegiorgis, B. Damene, R. Kakuhenzire, I. Kasahija, and I. Kahiu. 2011. Incentives and disincentives for stakeholder involvement in participatory research (PR): Lessons from potato-relateild PR from Bolivia, Ethiopia, Peru and Uganda. *International Journal of Agricultural Sustainability* 9: 522–536.

Pacini, G., D. Colucci, F. Baudron, E. Righi, M. Corbeels, P. Tittonell, and F. Stefanini. 2014. Combining multi-dimensional scaling and cluster analysis to describe the diversity of rural households. *Experimental Agriculture* 50: 376–397.

Pannell, D. J., G. R. Marshall, N. Barr, A. Curtis, F. Vanclay, and R. Wilkinson. 2006. Understanding and promoting adoption of conservation practices by rural landholders. *Australian Journal of Experimental Agriculture* 46: 1407–1424.

Pant, L. P., and H. Hambly Odame. 2009. The promise of positive deviants: Bridging divides between scientific research and local practices in smallholder agriculture. *Knowledge Management for Development Journal* 5: 160–172.

Patel, N., D. Chittamuru, A. Jain, P. Dave, and T. S. Parikh. 2010. Avaaj Otalo — A field atudy of an interactive voice forum for small farmers in rural India. CHI 2010: HCI and India, 733–742. Atlanta, USA: ACM.

Patel, N., K. Savani, S. R. Klemmer, and T. S. Parikh. 2013. Power to the peers: Authority of source effects for a voice-based agricultural information service in rural India. *Information Technologies & International Development* 9: 81–93.

Patnaik, S., E. Brunskill, and W. Thies. 2009. Evaluating the accuracy of data collection on mobile phones: A study of forms, SMS, and voice. Proceedings of the 2009 International Conference on Information and Communication Technologies and Development (ICTD). IEEE.

Perfect, J., and A. E. Majule. 2010. Livelihood zones analysis. A tool for planning agricultural water management investments: Tanzania. International Water Management Institute (IWMI).

Pitula, K., and T. Radhakrishnan. 2011. On eliciting requirements from end-users in the ICT4D domain. *Requirements Engineering* 16: 323–351.

Pretty, J. 2018. Intensification for redesigned and sustainable agricultural systems. *Science* 362: eaav0294.

Pretty, J., W. J. Sutherland, J. Ashby, J. Auburn, D. Baulcombe, M. Bell, J. Bentley, S. Bickersteth, K. Brown, J. Burke, H. Campbell, K. Chen, E. Crowley, I. Crute, D. Dobbelaere, G. Edwards-Jones, F. Funes-Monzote, H. C. J. Godfray, M. Griffon, P. Gypmantisiri, L. Haddad, S. Halavatau, H. Herren, M. Holderness, A.-M. Izac, M. Jones, P. Koohafkan, R. Lal, T. Lang, J. McNeely, A. Mueller, N. Nisbett, A. Noble, P. Pingali, Y. Pinto, R. Rabbinge, N. H. Ravindranath, A. Rola, N. Roling, C. Sage, W. Settle, J. M. Sha, L. Shiming, T. Simons, P. Smith, K. Strzepeck, H. Swaine, E. Terry, T. P. Tomich, C. Toulmin, E. Trigo, S. Twomlow, J. K. Vis, J. Wilson, and S. Pilgrim. 2010. The top 100 questions of importance to the future of global agriculture. *International Journal of Agricultural Sustainability* 8: 219–236.

Pretty, J., C. Toulmin, and S. Williams. 2011. Sustainable intensification in African agriculture. *International Journal of Agricultural Sustainability* 9: 5–24.

Qiang, C. Z., S. C. Kuek, A. Dymond, and S. Esselaar. 2012. Mobile applications for agriculture and rural development. ICT Sector Unit, Report No. 96226-GLB. Washington, D.C., USA: The World Bank.

Rejesus, R. M., M. E. M. Mutuc, M. Yasar, A. V. Lapitan, F. G. Palis, and T. T. N. Chi. 2012. Sending Vietnamese rice farmers back to school: Further evidence on the impacts of Farmer Field Schools. *Canadian Journal of Agricultural Economics* 60: 407–426.

Ricciardi, V., N. Ramankutty, Z. Mehrabi, L. Jarvis, and B. Chookolingo. 2018. How much of the world's food do smallholders produce? *Global Food Security* 17: 64–72.

Richards, P. 1989. Agriculture as a performance. *In:* R. Chambers, A. Pacey, and L. A. Thrupp (eds.). Farmer First: Farmer innovation and agricultural research. Intermediate Technology Publications.

Richards, P. 1993. Cultivation: Knowledge or performance? *In:* M. Hobart (ed.). An Anthropological Critique of Development. The Growth of Ignorance. London: Routledge.

Ritzema, R., R. Frelat, S. Douxchamps, S. Silvestri, M. C. Rufino, M. Herrero, K. E. Giller, S. López-Ridaura, N. Teufel, B. K. Paul, and M. T. van Wijk. 2017. Is production intensification likely to make farm households food-adequate? A simple food availability analysis across smallholder farming systems from East and West Africa. *Food Security* 9: 115–131.

Rosenstock, T. S., C. Lamanna, S. Chesterman, J. Hammond, S. Kadiyala, E. Luedeling, K. Shepherd, B. DeRenzi, and Mark T. van Wijk. 2017. When less is more: Innovations for tracking progress toward global targets. *Current Opinion in Environmental Sustainability* 26–27: 54–61.

Rufino, M. C., J. Dury, P. Tittonell, M. T. van Wijk, M. Herrero, S. Zingore, P. Mapfumo, and K. E. Giller. 2011. Competing use of organic resources, village-level interactions between farm types and climate variability in a communal area of NE Zimbabwe. *Agricultural Systems* 104: 175–190.

Ryan, S. F., N. L. Adamson, A. Aktipis, L. K. Andersen, R. Austin, L. Barnes, M. R. Beasley, K. D. Bedell, S. Briggs, B. Chapman, C. B. Cooper, J. O. Corn, N. G. Creamer, J. A. Delborne, P. Domenico, E. Driscoll, J. Goodwin, A. Hjarding, J. M. Hulbert, S. Isard, M. G. Just, K. Kar Gupta, M. M. López-Uribe, J. O'Sullivan, E. A. Landis, A. A. Madden, E. A. McKenney, L. M. Nichols, B. J. Reading, S. Russell, N. Sengupta, L. R. Shapiro, L. K. Shell, J. K. Sheard, D. D. Shoemaker, D. M. Sorger, C. Starling, S. Thakur, R. R. Vatsavai, M. Weinstein, P. Winfrey, and R. R. Dunn. 2018. The role of citizen science in addressing grand challenges in food and agriculture research. *Proceedings of the Royal Society B: Biological Sciences* 285: 20181977.

Saito, K., S. Diack, I. Dieng, and M. K. N'Diaye. 2015. On-farm testing of a nutrient management decision-support tool for rice in the Senegal River valley. *Computers and Electronics in Agriculture* 116: 36–44.

Schindler, J., F. Graef, and H. J. König. 2016. Participatory impact assessment: Bridging the gap between scientists' theory and farmers' practice. *Agricultural Systems* 148: 38–43.

Schlenker, W., and D. B. Lobell. 2010. Robust negative impacts of climate change on African agriculture. *Environmental Research Letters* 5: 014010.

Schut, M., L. Klerkx, M. Sartas, D. Lamers, M. McCampbell, I. Ogbonna, P. Kaushik, K. Atta-Krah, and C. Leeuwis. 2016. Innovation platforms: Experiences with their institutional embedding in agricultural research for development. *Experimental Agriculture* 52: 537–561.

Schut, M., J.-J. Cadilhon, M. Misiko, and I. Dror. 2018. Do mature innovation platforms make a difference in agricultural research for development? A meta-analysis of case studies. *Experimental Agriculture* 54: 96–119.

Schut, M., J. Kamanda, A. Gramzow, T. Dubois, D. Stoian, J. A. Andersson, I. Dror, M. Sartas, R. Mur, S. Kassam, H. Brouwer, A. Devaux, C. Velasco, R. J. Flor, M. Gummert, D. Buizer, C. McDougall, K. Davis, S. Homann-Kee Tui, and M. Lundy. 2019. Innovation platforms in agricultural research for development. Ex-ante appraisal of the purpose and conditions under which innovation platforms can contribute to agricultural development outcomes. *Experimental Agriculture* 55: 575-596.

Sheahan, M., and C. B. Barrett. 2017. Ten striking facts about agricultural input use in Sub-Saharan Africa. *Food Policy* 67: 12–25.

Snyder, K. A., and B. Cullen. 2014. Implications of sustainable agricultural intensification for family farming in Africa : Anthropological perspectives. *Anthropological Perspectives* 20: 9–29.

Sperling, L., J. A. Ashby, M. E. Smith, E. Weltzien, and S. McGuire. 2001. A framework for analyzing participatory plant breeding approaches and results. *Euphytica* 122: 106–119.

Spielman, D. J., J. Ekboir, and K. Davis. 2009. The art and science of innovation systems inquiry: Applications to Sub-Saharan African agriculture. *Technology in Society* 31: 399–405.

Steinke, J., and J. van Etten. 2016. Farmer experimentation for climate adaptation with triadic comparisons of technologies (tricot). A methodological guide. Rome, Italy: Bioversity International.

Steinke, J., J. van Etten, and P. Mejía Zelan. 2017. The accuracy of farmer-generated data in an agricultural citizen science methodology. *Agronomy for Sustainable Development* 37: 32.

Steinke, J., M. G. Mgimiloko, F. Graef, J. Hammond, M. T. van Wijk, and J. van Etten. 2019. Prioritizing options for multi-objective agricultural development through the Positive Deviance approach. *PLoS ONE* 14: e0212926.

Sternin, M., J. Sternin, and D. Marsh. 1998. Designing a community-based nutrition program using the Hearth model and the Positive Deviance approach – A field guide. Westport, USA: Save the Children.

Sulaiman, R., A. Hall, N. J. Kalaivani, and K. Dorai. 2012. Necessary, but not sufficient: Critiquing the role of information and communication technology in putting knowledge into use. *Journal of Agricultural Education and Extension* 18: 331–346.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* 142: 2282–2292.

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, D. Fink, J. W. Fitzpatrick, T. Fredericks, J. Gerbracht, C. Gomes, W. M. Hochachka, M. J. Iliff, C. Lagoze, F. A. La Sorte, M. Merrifield, W. Morris, T. B. Phillips, M. Reynolds, A. D. Rodewald, K. V. Rosenberg, N. M. Trautmann, A. Wiggins, D. W. Winkler, W.-K. Wong, C. L. Wood, J. Yu, and S. Kelling. 2014. The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation* 169.: 31–40.

Šūmane, S., I. Kunda, K. Knickel, A. Strauss, T. Tisenkopfs, I. des los Rios, M. Rivera, T. Chebach, and A. Ashkenazy. 2018. Local and farmers' knowledge matters! How integrating informal and formal knowledge enhances sustainable and resilient agriculture. *Journal of Rural Studies* 59: 232–241.

Sumberg, J., C. Okali, and D. Reece. 2003. Agricultural research in the face of diversity, local knowledge and the participation imperative: Theoretical considerations. *Agricultural Systems* 76: 739–753.

Suppe, F. 1987. The limited applicability of agricultural research. *Agriculture and Human Values* Fall 1987: 4–14.

Swindale, A., and P. Bilinsky. 2006. Household Dietary Diversity Score for measurement of household food access: Indicator guide. Washington, D.C., USA: Academy for Educational Development.

Tariq, A., A. De Neergaard, L. S. Stoumann, B. O. Sander, M. Van Trinh, Q. Duong Vu, R. Wassmann, and S. de Tourdonnet. 2018. Co-design and assessment of mitigation practices in rice production systems : A case study in northern Vietnam. *Agricultural Systems* 167: 72–82.

Tata, J. S., and P. E. McNamara. 2016. Social factors that influence use of ICT in agricultural extension in Southern Africa. *Agriculture* 6: 15.

Taylor, M., and S. Bhasme. 2018. Model farmers, extension networks and the politics of agricultural knowledge transfer. *Journal of Rural Studies* 64: 1–10.

Tittonell, P., M. T. van Wijk, M. C. Rufino, J. A. Vrugt, and K. E. Giller. 2007a. Analysing trade-offs in resource and labour allocation by smallholder farmers using inverse modelling techniques: A case-study from Kakamega district, western Kenya. *Agricultural Systems* 95: 76–95.

Tittonell, P., B. Vanlauwe, N. de Ridder, and K. E. Giller. 2007b. Heterogeneity of crop productivity and resource use efficiency within smallholder Kenyan farms : Soil fertility gradients or management intensity gradients? *Agricultural Systems* 94: 376–399.

Tittonell, P., A. Muriuki, K. D. Shepherd, D. Mugendi, K. C. Kaizzi, J. Okeyo, L. Verchot, R. Coe, and B. Valauwe. 2010. The diversity of rural livelihoods and their influence on soil fertility in agricultural systems of East Africa - A typology of smallholder farms. *Agricultural Systems* 103: 83–97.

Vanlauwe, B., D. Coyne, J. Gockowski, S. Hauser, J. Huising, C. Masso, G. Nziguheba, M. Schut, and P. van Asten. 2014. Sustainable intensification and the African smallholder farmer. *Current Opinion in Environmental Sustainability* 8: 15–22.

Vayssières, J., M. Vigne, V. Alary, and P. Lecomte. 2011. Integrated participatory modelling of actual farms to support policy making on sustainable intensification. *Agricultural Systems* 104: 146–161.

Verkaart, S., B. G. Munyua, K. Mausch, and J. D. Michler. 2017. Welfare impacts of improved chickpea adoption: A pathway for rural development in Ethiopia? *Food Policy* 66: 50–61.

Verkaart, S., K. Mausch, and D. Harris. 2018. Who are those people we call farmers ? Rural Kenyan aspirations and realities. *Development in Practice* 28: 468–479.

Wanvoeke, J., J. Venot, M. Zwarteveen, and C. de Fraiture. 2015. Performing the success of an innovation: The case of smallholder drip irrigation in Burkina Faso. *Water International* 40: 432–445.

van Wart, J., L. G. J. van Bussel, J. Wolf, R. Licker, P. Grassini, A. Nelson, H. Boogaard, J. Gerber, N. D. Mueller, L. Claessens, M. K. van Ittersum, and K. G. Cassman. 2013. Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Research* 143: 44–55.

Wichern, J., J. van Heerwaarden, S. de Bruin, K. Descheemaeker, P. J. A. van Asten, K. E. Giller, and M. T. van Wijk. 2018. Using household survey data to identify large-scale food security patterns across Uganda. *PLoS ONE* 13: e0208714.

Witcombe, B. J. R., A. Joshi, K. D. Joshi, and B. R. Sthapit. 1996. Farmer participatory crop improvement. I. Varietal selection and breeding methods and their impact on biodiversity. *Experimental Agriculture* 32: 445–460.

Witcombe, B. J. R., K. D. Joshi, S. Gyawali, A. M. Musa, C. Johansen, D. S. Virk, and B. R. Sthapit. 2005. Participatory plant breeding is better described as highly client-oriented plant breeding. I. Food indicators of client-orientation in plant breeding. *Experimental Agriculture* 41: 299–319.

Woodward, S. J. R., A. J. Romera, W. B. Beskow, and S. J. Lovatt. 2008. Better simulation modelling to support farming systems innovation: Review and synthesis. *New Zealand Journal of Agricultural Research* 51: 235–252.

Wossen, T., T. Abdoulaye, A. Alene, M. G. Haile, S. Feleke, A. Olanrewaju, and V. Manyong. 2017. Impacts of extension access and cooperative membership on technology adoption and household welfare. *Journal of Rural Studies* 54: 223–233.

Wright, H. J., W. Ochilo, A. Pearson, C. Finegold, M. L. Oronje, J. Wanjohi, R. Kamau, T. Holmes, and A. Rumsey. 2016. Using ICT to strengthen agricultural extension systems for plant health. *Journal of Agricultural and Food Information* 17: 23–36.

Wyche, S., and C. Steinfield. 2016. Why don’t farmers use cell phones to access market prices? Technology affordances and barriers to market information services adoption in rural Kenya. *Information Technology for Development* 22: 320–333.

9 Annex

Supplementary information

for:

Steinke, J., *et al.* (2019) Household-specific targeting of agricultural advice via mobile phones: Feasibility of a minimum data approach in smallholder context. *Computers and Electronics in Agriculture* 162:991-1000.

1. Procedure for determining information options used in ranking experiments

(i) Identification of locally viable, yet uncommon, household development options through the "Positive Deviance" approach (see Section 3.3)

To come up with a set of locally viable, yet not widely adopted practices to be ranked by respondents of the preference experiment, we implemented, at each site, the Positive Deviance approach (Steinke et al., 2019). This method consisted of two major steps: First, identifying individual households with surprisingly strong multi-dimensional performance from survey data ("positive deviant households"). Second, re-visiting these household for observation and identification of any uncommon, potentially innovative practices that plausibly explain their relative success ("positive deviant practices").

At each study site, we used the RHOMIS survey data (see Section 3.2) to calculate, for each household, five performance indicators as proxies for their achievement of five agreed key targets of sustainable intensification (Montpellier Panel, 2013; Smith et al., 2017; Snapp et al., 2018) (Table 1). Sustainable intensification is a widely established development paradigm for the farming population in the Global South (Godfray & Garnett 2014; Pretty et al., 2011; Rockström et al., 2017; Vanlauwe et al. 2014). At each study site, and for each of the five indicators, we then fit a median regression to data. These regressions included as covariates five known drivers of household performance: land endowments, livestock endowments, household size, region, and market access (Frelat at al., 2016). The regressions, thus, accounted for these external determinants of household performance, and helped to identify strong *relative* performance, irrespective of resources and location attributes. We then quantified each household's overall performance by its five residuals: A large residual indicates a large positive deviation from the expected performance. In this five-dimensional space of relative performance values, we then searched for Pareto-optimal solutions. Households with Pareto-optimal performance outperformed other households with equivalent characteristics in at least one of the five indicators, without being outperformed in another. It implies they achieved stronger overall multi-dimensional performance than other households with given resources. In the analysis, we used the package *emoa* (Mersmann, 2012) in the R software environment (R Core Team, 2018).

Table 1: Agreed targets in sustainable intensification and indicators used to approximate household performance per target. Performance indicator definitions are given below.

Development target	Household performance indicator
Food security	Caloric food security
Healthy nutrition	Dietary diversity
Income	Cash income
Environmental sustainability	Greenhouse gas emissions
Social equity	Gender equity

At each site, we then re-visited a diverse sample of these "positive deviant" households to identify the specific practices that were already being successfully implemented by the positive deviants as strategies to achieve superior multi-dimensional performance. Through in-depth interviews and farm observations, we intended to document any uncommon practices that plausibly explained their relative success. Through visits to 14 positive deviants in Ethiopia, we identified 14 practices. With 10 positive deviants in Kenya, we identified 9 practices. With 15 positive deviants in Tanzania, we identified 14 practices. At each site, we selected 9 practices as information options for our experiment. In Kenya, this was the total number of practices found with positive deviants. In Ethiopia and Tanzania, we chose practices that could realistically be supported by the local agricultural advisory services through information interventions.

(ii) To Table 1 (above): Performance indicator definitions used to define multi-dimensional household performance

Development target: **Food security**

Performance indicator: **Caloric food security**

Defined by first loading of a principal component analysis on two key indicators of household food security:

- *Sufficiency of access to food.* We estimated household calorie needs by multiplying household size (in male adult equivalent, MAE) by 2,550 Kcal, Tanzania's official recommended daily energy intake per MAE (Perfect and Majule 2010). Then, household food availability, i.e. the potential amount of caloric food energy generated by all on- and off-farm activities of the household (from RHoMIS, see Hammond et al. 2017 and Frelat 2016), was divided by the obtained value, and capped at 100 %.
- *Consistency of access to food.* Number of months that are considered to be food-secure on a 12 month recall period.

Development target: **Nutrition**

Performance indicator: **Household Diet Diversity Score**

Harmonic mean of the HDDS values in the (self-defined) good season and lean season. HDDS is defined as the number of items out of 12 food groups (e.g., starch crops, legumes, green vegetables, eggs, meat, etc.) that are consumed by the household at least once per week (Swindale and Bilinsky 2006) during the latest recall season.

Development target: **Income**

Performance indicator: **Cash income**

Sum of incomes from farmgate sales and off-farm sources during the 12 months recall period.

Development target: **Environmental sustainability**

Performance indicator: **Greenhouse gas emissions**

Sum of CO_2-equivalents emitted by reported household activities. Calculated from standard emission values from literature using the IPCC Tier 1 approach (IPCC 2006). We then multiplied values by -1 because higher emissions imply weaker sustainability.

Development target: **Social equity**

Performance indicator: **Gender equity**

Defined through the gendered distribution of agriculture- and food-related decision-making. For the three domains of production decisions, consumption decisions, and marketing decisions, women's shares of decision-making were calculated and averaged. Different household types (woman-led, man-led, single woman, etc.) were accounted for.

2. Explanations of practices included as information options in preference ranking experiments

Information option	Code	Explanation
Ethiopia		
Sowing cereals in lines	L	Applying recommended crop-specific spacing recommendations rather than broadcasting
Diligent farm scheduling	S	Informed, careful timing and coordination of different field operations
Rain water harvesting	R	Construction of rain water catchment ponds for collective or individual irrigation use
Storing and trading crops	T	Engaging in buying, bulking, transport and sale of crops
Opening a business	B	Generating income through self-employed sale of non-agricultural goods and services, such as drinks or hair styling
Tree nursery	N	Producing tree seedlings for sale
Reducing food wastage	W	Applying strategies that reduce the share of food that is cooked, but not eaten and thus wasted
Finding off-farm job	J	Generating income through off-farm wage labour
Improving crop storage	C	Decreasing post-harvest losses, e.g. by investing into improved crop storage constructions or triple layer "PICS" sacks
Kenya		
Machine tillage	M	Renting a fuel-driven tillage device or commissioning a tractor-tillage provider
Terracing	T	Construction of terraces for erosion control and improved soil moisture management

Zai pits	Z	Construction of zait pits for improved soil moisture management
Dry planting	D	Increasing the growing period by sowing and planting before the onset of rains, when soil is still hard
Collective crop marketing	G	Pooling produce and bargaining power within a group of smallholder producers
Mulching	R	Applying plant residues to the soil for erosion control and improved soil moisture management
Opening a business	B	Generating income through self-employed sale of non-agricultural goods and services, such as drinks or hair styling
Renting out traction animals	O	Renting own oxen to other farm households for tillage operations
Finding off-farm job	J	Generating income through off-farm wage labour
Tanzania		
Intercropping Pigeon pea / Maize	I	Implementing a specific, particularly resource-efficient intercropping method
Improving crop storage	C	Decreasing post-harvest losses, e.g. by investing into improved crop storage constructions or triple layer "PICS" sacks
Diligent farm scheduling	S	Informed, careful timing and coordination of different field operations
Machine tillage	M	Renting a fuel-driven tillage device or commissioning a tractor-tillage provider
Intensifying poultry production	P	Increasing poultry production through improved breeds and more secure coops
Opening a business	B	Generating income through self-employed sale of non-agricultural goods and services, such as drinks or hair styling
Tree nursery	T	Producing tree seedlings for sale

"Livestock bank"	L	Maintaining ruminant livestock against short-term utility logic, for sale in emergency situations
Finding off-farm job	J	Generating income through off-farm wage labour

References

Frelat, R., S. Lopez-Ridaura, K.E. Giller, M. Herrero, S. Douxchamps, A. Andersson Djurfeldt, O. Erenstein, B. Henderson, M. Kassie, B.K. Paul, C. Rigolot, R.S. Ritzema, D. Rodriguez, P.J.A. van Asten, M.T. van Wijk, 2016. Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *PNAS* 113, 458–463.

Godfray, H.C.J., T. Garnett, 2014. Food security and sustainable intensification. *Philosophical Transactions of the Royal Society B* 369: 20120273.

Hammond, J., S. Fraval, J. van Etten, J. Gabriel, L. Mercado, T. Pagella, R. Frelat, M. Lannerstad, S. Douxchamps, N. Teufel, D. Valbuena, M.T. van Wijk, 2017. The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. *Agricultural Systems* 151, 225–233.

IPCC, 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html

Mersmann, O., 2012. emoa: Evolutionary Multiobjective Optimization Algorithms. R package version 0.5-0.

Montpellier Panel, 2013. Sustainable intensification: A new paradigm for African agriculture. London.

Perfect, J., A.E. Majule, 2010. Livelihood zones analysis. A tool for planning agricultural water management investments: Tanzania. International Water Management Institute (IWMI).

Pretty, J., C. Toulmin, S. Williams, 2011. Sustainable intensification in African agriculture. *International Journal of Agricultural Sustainability* 9, 5–24.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rockström, J., J. Williams, G. Daily, A. Noble, N. Matthews, L. Gordon, H. Wetterstrand, F. Declerck, M. Shah, P. Steduto, C. de Fraiture, N. Hatibu, O. Unver, J. Bird, L. Sibanda, J. Smith, 2017. Sustainable intensification of agriculture for human prosperity and global sustainability. *Ambio* 46, 4–17.

Smith, A., S. Snapp, R. Chikowo, P. Thorne, M. Bekunda, J. Glover, 2017. Measuring sustainable intensification in smallholder agroecosystems: A review. *Global Food Security* 12, 127–138.

Snapp, S.S., P. Grabowski, R. Chikowo, A. Smith, E. Anders, D. Sirrine, V. Chimonyo, M. Bekunda, 2018. Maize yield and profitability tradeoffs with social, human and environmental performance: Is sustainable intensification feasible? *Agricultural Systems* 162, 77–88.

Steinke, J., M.G. Mgimiloko, F. Graef, J. Hammond, M.T. van Wijk, J. van Etten, 2019. Prioritizing options for multi-objective agricultural development through the Positive Deviance approach. *PLoS One* 14(2):e0212926.

Swindale, A., P. Bilinsky, 2006. Household Dietary Diversity Score for measurement of household food access: Indicator guide. FHI 360/FANTA, Washington, D.C.

Vanlauwe, B., D. Coyne, J. Gockowski, S. Hauser, J. Huising, C. Masso, G. Nziguheba, M. Schut, P. van Asten, 2014. Sustainable intensification and the African smallholder farmer. *Current Opinion in Environmental Sustainability* 8, 15–22.

Eigenständigkeitserklärung

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegegebenen Hilfen und Hilfsmittel angefertigt zu haben.

Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad. Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde. Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 5. März 2015.

Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen PromotionsberaterInnen stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftliche Praxis eingehalten wurden.

Statutory declaration

I hereby declare having completed the doctoral thesis independently and solely based on the stated resources and aids.

I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree. I have not submitted the doctoral thesis, or parts of it, to another academic institution and the thesis has not been accepted or rejected. I declare having acknowledged the Doctoral Degree Regulations of the Faculty of Life Sciences at Humboldt-Universität zu Berlin of March 5, 2015, which underlie the procedure.

Furthermore, I declare no collaboration with any commercial doctoral degree supervisors took place, and that the principles of Humboldt-Universität zu Berlin for ensuring good academic practice were abided by.

Berlin, 22 August 2019

Jonathan Steinke